GU J Sci 30(4): 373-393 (2017)

Gazi University

Journal of Science

http://dergipark.gov.tr/gujs

Graph Communities To Analyze The Occupational Accidents: An Evidence From The Statistics Of Turkey 2013–2014

Gökhan TUNA[1,*], Mustafa KURT[2]

[1]*Gazi University, Graduate School of Natural and Applied Sciences, Department of Environmental and Technical Research of Accidents, 06500, Ankara, Turkey*

[2]*Gazi University, Faculty of Engineering, Department of Industrial Engineering, 06500, Ankara, Turkey*

Article Info

Received: 31/03/2017
Accepted: 12/10/2017

Keywords

Occupational Accidents
Network Analysis
Pearson Correlation
Graph Communities
Minimum Spanning Tree

Abstract

In this study, we present an efficient graph clustering method called graph communities to analyze occupational accidents in Turkey from 2013 to 2014 and use the Pearson correlation to calculate the similarity measure between the accident data. For this purpose we represent each accident as a vector of a social actor. Then, we obtain sub-communities which can be seen as the cluster of clusters iteratively. Moreover, we present the sub-dominant ultra-metric structures among occupational accidents using the minimum spanning tree method. In each sub-community we analyze the central nodes in minimum spanning trees which are the dominant in the information flow. Furthermore, we use the topological measures to determine which hierarchical structure is best to represent occupational accidents in sub-communities.

1. INTRODUCTION

The increase in the number of employees and the diversity of the business increases the material and moral damages to workers, their families and the economic dynamics of a country caused by the occupational accidents. Occupational accidents are happening depended on several reasons such as manufacturing type, manufacturing tools, working environments, and psychological factors.

Institutions attach importance to occupational health and safety for legal and risk analysis reasons [9]. Traditional and descriptive methods are the most common ones. The increase in the number of employees and the diversity of the business increases the material and moral damages to workers, their families and the economic dynamics of a country caused by the occupational accidents. Occupational accidents are happening depended on several reasons such as manufacturing type, manufacturing tools, working environments, and psychological factors.

Institutions attach importance to occupational health and safety for legal and risk analysis reasons [9]. Traditional and descriptive methods are the most common ones to analyse occupational accidents. These methods mainly consider the chronological order of the collection of the information related to the occupational accident situation [8,12,18,23,24,39]. However, there are few studies on the correlation of the type of the accident and all factors leading and impacting the accident. In [8], authors briefly show that junior workers and the workers who have less information about the workplace are more tend to have occupational accidents. The mutual safety zone of the working group is considered as the accident prediction factor in [12]. Factors of occupational accidents in construction areas are explained as hardware inadequacy, communication problems between working groups, workplace problems, risk maintenance problems and inappropriateness of the tools in [18]. In [23], authors study the problems of

*Corresponding author, e-mail: gokhantuna@hotmail.com

insomnia in several economic activity groups. Generally, these studies focus on a certain kind of problem or economic activity group that cause the leading occupational accidents. Hence, these traditional and descriptive methods cannot reach more detailed results and they remain as only a chronological order study.

Examination of occupational accidents is important in the prevention of new accidents of the same type in the workplace and in the discovery of incorrect and problematic issues. However, it is also important in terms of deterrence that the faulty and responsible persons related to the accident are detected and punished. Work accidents are complex events that have led to study on how these accidents happen. A scientific change in questioning the causes of work accidents have occurred in the early 1930s [19,33]. The evolution of the methods starts from the simple domino models that examine individuals' behaviour and more sophisticated linear models that deal with the time series of event analysis into systematic models that deal with obstacles and defences. Newer and more complex and non-linear models have emerged with greater attention to the complexity of the causes of work accidents. Some important ones are Gordon's Multi Factor Model [15], Haddon's Energy Transfer Model [17], Wigglesworth's Injury Causation Model [48], and Peterson's Humanistic Model [37].

Many methods based on time series analysis are used for the analysis of work accidents and for the measures to be taken after the determination of the causes. However, in the literature, apart from statistical methods, it is seen that methods such as graph theory, which are very useful methods in modelling complex systems, are not used. Successful use of mathematical methods is increasingly being used to account for phenomena observed in the social, economic, and biological systems of the real world [2,3,4,6,40]. In [16], authors present a meaningful demonstration of the complex relations between the units in detail on many different systems. The numerical results obtained as a result of these visualization methods provide positive contributions to social systems in terms of research. Hierarchical methods such as graph theory have a more transparent structure than quantitative methods such as data mining and statistical regression. Recently; Complex system tools such as communication networks and hierarchical notations used in the fields of economy, sociology and medicine have accelerated the research in these fields. However, as long as real data is not available, such system models are not exact [29].

Complex systems consisting of too many non-linearly related agents are natural or social systems. Researchers from different institutions that are working on new models and trying to apply complex methods from different disciplines have heard the need to collect data to understand phenomenology with an easier analysis. It is the most eye-catching feature of these systems to obtain phenomena that are not predicted clearly or are not derived easily by plain information of the system's structure. Graph theory methods have been developed between many complex system properties, such as long-term interactions [36], measure invariance [25], aggregation effects and noise [11], criticality [31], and the determinism and flexibility in the evolution [20].

The social network analysis of resource change and resource appropriation among individuals examines and focuses on the relationship model of the individuals in the interaction [21,38,46]. Abstractly, social support and influence items, in concrete terms, such as goods, services and money, can be the source of resource change. With social network analysis, the distributions of these abstract items in the social networks of workers can be examined. There is a relationship that corresponds to a particular type of abstract item change. Individuals or organizations may be actors who change these items. It specifies information associations about which type of information has been changed among whom. It is the model of the relationship between the actors who take the authoritarian form and bring out the similarities of the extraction of certain types of information to the individual. These models explain how actors control the flow of information and how information is circulated in networks.

Social network analysis, which analyses models that deal with issues such as working relationship or information traffic, is more advanced than other analysis methods. Therefore, social network analysis will be adopted as an effective method to analyse each worker's accident as an actor and to analyse how the exchange of information between actors or abstract element change occurs. The empirical technique of finding the observed relationship between actors is a social network analysis technique that uses social construction as opposed to a priori classification. The classification of actors' titles, roles, ages, genders, and status according to similarities are typical non-networking approaches.

Researchers define classes by assigning individuals to classes according to "the mental method of comparing where similar objects are grouped together as a concept" [7]. The common label in this method is the "group" definition. By examining the interaction between the groups at the next stage, the researchers visualize the social structure. Social network analysis techniques examine the actors before they group related actors. Groups are formed in a network of closely related regions [30]. Other actors within the same classification will need to be labelled later as a group in order to predict their behaviours and make meaningful configuration [7]. The usefulness of the category of a group arises from the concentration of the actor's interactions with each other, not from the membership of a particular ethnic origin, gender, class, etc.

Complex systems in which occupational accidents are treated as actors can also be visualized with graphs. The complexity of such a system will also be shown on the related graph of the system. This graph is with a structure of intense connections and complicated relations. For this reason, it will be possible to switch to a less complicated sub-structure of such a complex graph by using a number of filtering methods to determine its important properties. Minimum Spanning Tree (MST) is one of the most commonly used filtering methods in the literature. The spanning three of a graph G is an acyclic subgraph that contains all the vertices of G. MST algorithms traditionally operate with $O(mlogn)$ complexity [10], where m is the number of the edges in the graph, and n is the number of graph vertices. The Correlation Minimum Spanning Tree (CMST) is the spanning tree of points in the n-dimensional Euclidean space with respect to the correlation of each vertices. The correlation distance and edge lengths of two points of this tree are equal [28]. Specifying clusters with irregular bounds is the capability of the MST clustering algorithm. The MST clustering algorithms differ from traditional clustering algorithms in a way that they do not consider the global shape clustering scheme.

The MST, which selects the most relevant links for each point in the data, is the most preferred approach. Furthermore, this algorithm directly gives an ultra-metric sub-dominant hierarchical structure of the points of the data set analysed [28]. The isolated groups of actors in the complex system that we have identified as work accidents can be identified by considering the topology that emerges from the relevant hierarchical structure found by the MST. The identified hierarchical structure is searching for factors that affect isolated groups and will be effective in the theoretical definition of work accidents. The taxonomy of this hierarchical structure will be found by the parametric inputs of the work accident data.

In [5], author shows that graph communities are effective clustering methods in complex systems. While reaching the hierarchical structure, firstly, a weighted complete graph of the actors is taken into consideration. A single community on this weighted graph will appear to include all actors. In order to come up with this problem, an algorithmic way of generating graphs to include the most appropriate number of links depending on the eigenvalues of the graph modelling the complex system has been followed. The hierarchies in the formed communities are formed strongly in terms of both taxonomic and clusters. When this method is applied on the data structures determined by parametric inputs, like the data we provide in our study, the correlation distances are very close to zero, that is, they are highly correlated and will not give effective results in complex structures. In order to overcome this problem, a similar method to [5] is developed in our study and a hierarchical structure analysis has been made by examining the data of the occupational accidents in Turkey in 2013–2014.to analyze occupational accidents. These methods mainly consider the chronological order of the collection of the information related to the occupational accident situation [8,12,18,23,24,39]. However, there are few studies on the correlation of the type of the accident and all factors leading and impacting the accident. In [8], authors briefly show that junior workers and the workers who have less information about the workplace are more tend to have occupational accidents. The mutual safety zone of the working group is considered as the accident prediction factor in [12]. Factors of occupational accidents in construction areas are explained as hardware inadequacy, communication problems between working groups, workplace problems, risk maintenance problems and inappropriateness of the tools in [18]. In [23], authors study the problems of insomnia in several economic activity groups. Generally, these studies focus on a certain kind of problem or economic activity group that cause the leading occupational accidents. Hence, these traditional and descriptive methods cannot reach more detailed results and they remain as only a chronological order study.

Examination of occupational accidents is important in the prevention of new accidents of the same type in the workplace and in the discovery of incorrect and problematic issues. However, it is also important in terms of deterrence that the faulty and responsible persons related to the accident are detected and punished. Work accidents are complex events that have led to study on how these accidents happen. A scientific change in questioning the causes of work accidents have occurred in the early 1930s [19,33]. The evolution of the methods starts from the simple domino models that examine individuals' behaviour and more sophisticated linear models that deal with the time series of event analysis into systematic models that deal with obstacles and defences. Newer and more complex and non-linear models have emerged with greater attention to the complexity of the causes of work accidents. Some important ones are Gordon's Multi Factor Model [15], Haddon's Energy Transfer Model [17], Wigglesworth's Injury Causation Model [48], and Peterson's Humanistic Model [37].

Many methods based on time series analysis are used for the analysis of work accidents and for the measures to be taken after the determination of the causes. However, in the literature, apart from statistical methods, it is seen that methods such as graph theory, which are very useful methods in modelling complex systems, are not used. Successful use of mathematical methods is increasingly being used to account for phenomena observed in the social, economic, and biological systems of the real world [2,3,4,6,40]. In [16], authors present a meaningful demonstration of the complex relations between the units in detail on many different systems. The numerical results obtained as a result of these visualization methods provide positive contributions to social systems in terms of research. Hierarchical methods such as graph theory have a more transparent structure than quantitative methods such as data mining and statistical regression. Recently; Complex system tools such as communication networks and hierarchical notations used in the fields of economy, sociology and medicine have accelerated the research in these fields. However, as long as real data is not available, such system models are not exact [29].

Complex systems consisting of too many non-linearly related agents are natural or social systems. Researchers from different institutions that are working on new models and trying to apply complex methods from different disciplines have heard the need to collect data to understand phenomenology with an easier analysis. It is the most eye-catching feature of these systems to obtain phenomena that are not predicted clearly or are not derived easily by plain information of the system's structure. Graph theory methods have been developed between many complex system properties, such as long-term interactions [36], measure invariance [25], aggregation effects and noise [11], criticality [31], and the determinism and flexibility in the evolution [20].

The social network analysis of resource change and resource appropriation among individuals examines and focuses on the relationship model of the individuals in the interaction [21,38,46]. Abstractly, social support and influence items, in concrete terms, such as goods, services and money, can be the source of resource change. With social network analysis, the distributions of these abstract items in the social networks of workers can be examined. There is a relationship that corresponds to a particular type of abstract item change. Individuals or organizations may be actors who change these items. It specifies information associations about which type of information has been changed among whom. It is the model of the relationship between the actors who take the authoritarian form and bring out the similarities of the extraction of certain types of information to the individual. These models explain how actors control the flow of information and how information is circulated in networks.

Social network analysis, which analyses models that deal with issues such as working relationship or information traffic, is more advanced than other analysis methods. Therefore, social network analysis will be adopted as an effective method to analyses each worker's accident as an actor and to analyze how the exchange of information between actors or abstract element change occurs. The empirical technique of finding the observed relationship between actors is a social network analysis technique that uses social construction as opposed to a priori classification. The classification of actors' titles, roles, ages, genders, and status according to similarities are typical non-networking approaches.

Researchers define classes by assigning individuals to classes according to "the mental method of comparing where similar objects are grouped together as a concept" [7]. The common label in this method is the "group" definition. By examining the interaction between the groups at the next stage, the researchers visualize the social structure. Social network analysis techniques examine the actors before they group related actors. Groups are formed in a network of closely related regions [30]. Other actors within the same classification will need to be labelled later as a group in order to predict their behaviors and make meaningful configuration [7]. The usefulness of the category of a group arises from the concentration of the actor's interactions with each other, not from the membership of a particular ethnic origin, gender, class, etc.

Complex systems in which occupational accidents are treated as actors can also be visualized with graphs. The complexity of such a system will also be shown on the related graph of the system. This graph is with a structure of intense connections and complicated relations. For this reason, it will be possible to switch to a less complicated sub-structure of such a complex graph by using a number of filtering methods to determine its important properties. Minimum Spanning Tree (MST) is one of the most commonly used filtering methods in the literature. The spanning three of a graph G is an acyclic subgraph that contains all the vertices of G. MST algorithms traditionally operate with O(mlogn) complexity [10], where m is the number of the edges in the graph, and n is the number of graph vertices. The Correlation Minimum Spanning Tree (CMST) is the spanning tree of points in the n-dimensional Euclidean space with respect to the correlation of each vertex. The correlation distance and edge lengths of two points of this tree are equal [28]. Specifying clusters with irregular bounds is the capability of the MST clustering algorithm. The MST clustering algorithms differ from traditional clustering algorithms in a way that they do not consider the global shape clustering scheme.

The MST, which selects the most relevant links for each point in the data, is the most preferred approach. Furthermore, this algorithm directly gives an ultra-metric sub-dominant hierarchical structure of the points of the data set analyzed [28]. The isolated groups of actors in the complex system that we have identified as work accidents can be identified by considering the topology that emerges from the relevant hierarchical structure found by the MST. The identified hierarchical structure is searching for factors that affect isolated groups and will be effective in the theoretical definition of work accidents. The taxonomy of this hierarchical structure will be found by the parametric inputs of the work accident data.

In [5], author shows that graph communities are effective clustering methods in complex systems. While reaching the hierarchical structure, firstly, a weighted complete graph of the actors is taken into consideration. A single community on this weighted graph will appear to include all actors. In order to come up with this problem, an algorithmic way of generating graphs to include the most appropriate number of links depending on the eigenvalues of the graph modelling the complex system has been followed. The hierarchies in the formed communities are formed strongly in terms of both taxonomic and clusters. When this method is applied on the data structures determined by parametric inputs, like the data we provide in our study, the correlation distances are very close to zero, that is, they are highly correlated and will not give effective results in complex structures. In order to overcome this problem, a similar method to [5] is developed in our study and a hierarchical structure analysis has been made by examining the data of the occupational accidents in Turkey in 2013–2014.

2. METHODOLOGY

2.1. Graph Theory Preliminaries

Graphs are one of the most powerful mathematical tools to model interacting agents. Mathematically, a graph G can be expressed with a tuple $G = (V, E)$ where V is a non-empty set of nodes (or vertices) and E is the set of node pairs. The elements of E are called the links (or edges) and can be expressed with the pair $e_i = (u, v)$ for the nodes u and v. If there is a link between any nodes of V, then these nodes are called adjacent. Generally, a graph G is the illustration of a relation defined on $V \times V$. A non-negative value called weight can be assigned to each links in this relation. The graph with the weight function $\omega: E \rightarrow \mathbb{R}^+$ is called a weighted graph and is expressed with the triple $G = (V, E, \omega)$. If (u, v) and (v, u) are representing the same link, that is the ordering of the relation is irrelevant, then G is called an undirected graph. Throughout this study we let $G = (V, E)$ be an undirected graph and not involve the loops $e_i = (u, u)$. The directed graph analogues of the definitions and results we use in this study can be found in [10].

The relation between the nodes of a graph can be expressed with a binary matrix $A(G)$ whose entries are

$$a_{ij} = \begin{cases} 1, & if\ (v_i, v_j) \in E \\ 0, & otherwise \end{cases}$$

where $V = \{v_1, \dots, v_n\}$. We shall note that $A(G)$ is symmetric by definition, thus all eigenvalues of $A(G)$ are real.

A degree of a node v is the number of the nodes adjacent to v. The communication can be considered as the sequence of links and if there is a communication between each nodes of G, then G is called a connected graph. The node is called isolated if there is no communication between those nodes to others. If all nodes of the graph are isolated then the graph is called a null graph, and if there is a link between all nodes of the graph then the graph is called a complete graph.

The Laplacian Matrix $L(G)$ of G is the $|V| \times |V|$ type matrix whose entries are

$$l_{ij} = \begin{cases} d_{v_i}, & if\ i = j \\ -1, & if\ A(G)(i,j) = 1 \\ 0, & otherwise \end{cases}$$

where d_{v_i} is the degree of v_i. The spectrum of $L(G)$ is positive semi-definite and the multiplicity k of the 0 eigenvalue equals the connected components in G [43]. In other words, if $L(G)$ is with only one 0 eigenvalue, then one can conclude that G involves no isolated nodes.

2.2. Network Construction and Communities

As the nodes represent interacting agents in a complex networks, links are issued to represent relations. There are several methods to determine link formation rules. For instance, in financial networks, the relation between different stock markets or exchange currencies is determined by the correlation or dynamic time warping distances [28,45]. Each node can be expressed by time series and then the distance between them leads the weight of the relation. In [5], author presents a novel approach to determine tuned topology of the sub-dominant ultra-metric structures in the data clusters. In this study we present a modified approach of the method presented in [5].

Rather than the time series, we first let the agents represented by the Euclidean N-vectors, that is $v_i = (v'_1, , \dots, v'_M) \subset \mathbb{R}^M$ for $i = 1, \dots, N$ where $|V\ | = N$. To form the links among the interacting agents we also consider the Pearson Correlation of each vector as

$$\rho_{ij} = \frac{E[v_i, v_j] - E[v_i]E[v_j]}{\sqrt{E[{v_i}^2] - (E[v_i])^2}\sqrt{E[v_j^2] - (E[v_j])^2}}.$$

Since ρ_{ij} varies in between -1 and 1, we consider weight of the links as the correlation distance $d_{corr}(v_i, v_j) = \sqrt{2(1\ -\ \rho_{ij}\ }) / 2$.

The correlation distance plays also a key role to determine the link formation rule with a threshold value. As suggested in [5], we also initially start with a complete graph of N-agents. 0 occurs with multiplicity of 1 in the spectrum of this initial graph. Then, by subdividing $[0,1]$ interval with $1/h$ step size we are able to obtain such step TV that the graph becomes with more than one component if we form the links with the rule $(v_i, v_j) \in E \;\; iff \; d_{corr}(v_i, v_j) \leq T\,V$.

Once we have an optimized many links in the network respect to threshold value, it becomes possible to figure out meaningful clusters via graph communities. Graph communities are the clusters of nodes that densely connected [13,26]. Graph community clustering aims to determine these nodes by using the topology of the network. To analyse graph communities can be efficient in social networks. For instance, a graph community could be common interest, location, or function of actors of the network. By determining communities, it is possible to study communities individually. Since different communities often exhibit different properties, individual ones let us to analyse actors meaningfully. Besides, by considering each community as "meta-node", it becomes possible to analyse networks with huge number of agents. In this study we use the High Modularity method to determine communities. The details of the method can be found in [1].

In the case of the relation of the complex network is determined by the respectful distance of the time series with long enough range, one may obtain several communities. However, in the case of vectors embedded lower dimension Euclidean space the correlation between these vectors would yield lesser communities. In this study we present a method to overcome this problem as considering coarser levels of the communities; i.e. we find communities of communities until we reach empirically useful clusters. Furthermore, if we fix TV that emerges from the first network then it is possible to construct subnetworks in the same sense since TV is a hereditary property for sub–topologies. In Figure 1, the coarser levels of the communities are shown for a regular partition. For the n-th level of the partition, one can obtain many 2^n communities. Each community involves the same TV of the ancestors Community 1 and Community 2.

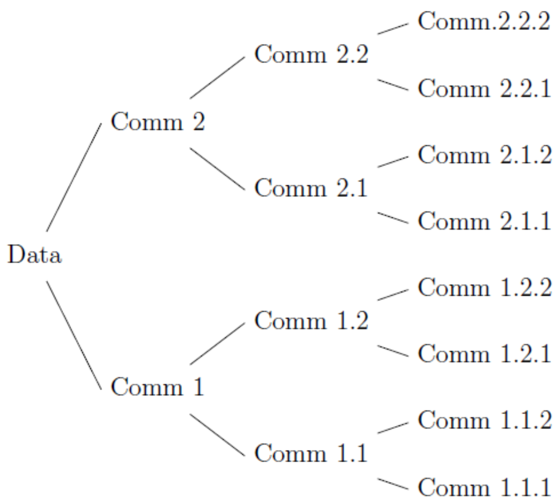

Figure 1. *The regular 2 communities' partition of the Data in 3 Levels*

"Comm" is the abbreviation for "Community".

2.3. Minimum Spanning Tree

One of the most efficient subgraphs to extract topological information of a complex network is minimum spanning tree (MST). The MST is a graph that connects all the n nodes of the graph with $(n-1)$ edges in such way that the total edge weights are a minimum. For the distance matrix of the complex network, MST reduces the information space from $n(n-1)/2$ correlations to $n-1$ tree edges. The MST of a complex network may not be unique. However, if the weights of all the edges are pairwise distinct, it is indeed unique [27]. The uniqueness of the MST implies that it is possible to compare MSTs of different weighted graphs when each graphs have the same node set, connected and have unique weights. This uniqueness is also important because one may not need to an arbitrary threshold value to reconstruct the graph.

MST has a simpler structure than the original graph; hence the analysis gets more simplified. However, due to this structure the tree also cannot reflect some properties which are dependent on the cycles. As presented in [41], this is not a disadvantage as it is assumed. For instance, under some conditions the flow of the information on the networks is dominated by the MST. Besides, the information flow in a weighted net-work is mainly limited to the MST in the strong disorder limit, with a high variability of link weights [45].

3. RESULTS

In this study, a social network of occupational accidents are modelled by a simple undirected graph $G = (\mathrm{V}, \mathrm{E})$, where V is the set of vectors of workers and E is the set of links that determined by the Pearson Correlation among the accidents. The data we used is obtained from the Statistics of Social Security Institution of Republic of Turkey (SGK) for the calendar years of 2013–2014.

3.1. Data

3.1.1. Restriction

An effective way to analyse the information on occupational accidents is to express the number of accidents in relation to the number of persons employed. This is referred as the incidence rate and gives an indication of the likelihood of someone having an accident [32,42]. General incidence rate or Type-I incidence rate value R_1 is calculated for the total fatal and non-fatal number of occupational accidents as

$$R_1 = \frac{Number\ of\ the\ accidents \times 1000}{Number\ of\ workers}$$

for per 1000 workers.

Occupational accidents in Turkey are reported to the SGK as a legal obligation. SGK compiles and reports the statistics obtained from these notifications. The classification of the activity groups in these reports uses the "Statistical Classification of Economic Activities in the European Community" (NACE). The last report published is for 2014. The average of general data and the incidence rates of work accidents (according to R_1) were calculated for the last 7 years (2014-2008) SGK statistical yearbooks. The R_1 values are given in Table 1.

At the occupational accident rate, 10 activity groups are above the country average. 179 (178.58887) out of 1000 workers working in "Coal and Lignite Removing", ranked 1st in terms of work accident incidence rates, reached the result of work accident. In the country average, 9.9 (9.8882) people have a work-related accident per 1000 people. In this study, the rate of incidence of occupational accidents constitutes the enforcement of 10 activity groups which are above the country average.

In this study, we restrict our data to the economic groups which have the incidence rate above the total of Turkey to catch more precise network.

Table 1. *Incidence Rates R_1 of Economic Activity Groups in Turkey during 2008–2014*

Economic Activity Group	R_1	NACE Code
Coal and Lignite Removing	178.5888	05
Main Metal Industry	47.3559	24
Non-Metal Goods Manufacturing	28.7251	23

Factory-made Metal Manufacturing	28.0157	25
Machine and Equipment Manufacturing	21.4211	28
Rubber and Plastic Goods Manufacturing	20.6713	22
Textile Goods Manufacturing	15.5078	13
Electric Gear Manufacturing	13.9254	27
Food Manufacturing	12.1880	10
Non-Building Construction	12.0597	42
Turkey Total	9.8882	
Storage and Support Operation for Logistics	9.8666	52
Special Construction Activities	9.7963	43
Food and Beverage Serving Activities	8.2147	56
Building Construction	7.1875	41
Land and Pipe Carriage	7.1716	49

3.1.2. Vectors

Agents in a complex network are often modelled as operating within social environment consisting of a network of interactions with other agents. Also, it is sometimes useful to model agents within a physical environment. This approach yields that agents that are farther apart are lesser able to influence each other. Hence, to consider "distance" between the agents, we represent them with vectors in N-dimensional space. By this approach, the physical environment of the agents becomes more like to be abstract and statistical operations can be more precisely applied.

The vectors to express each occupational accident in Turkey in the time span of 2013–2014 are with 13 components that are they are embedded in $\mathbb{R}^{13}$. The parameters to determine each component of the vector that is expressing the statistical state of the worker have had an occupational accident are Age, Gender, Technical Training, Occupational Safety Training, Working Year in the Last Job, Working Year in the Total, Environment of the Accident, Place of the Accident, Workplace, Environment of the Workplace, Cause of the Injury, Type of the Injury, Place of Injury in the Body, respectively.

The age of the workers are considered as the greatest integer return. The gender of the worker is considered as 1 for males and 2 for females. Technical and occupational safety training are considered as 1 for Yes and 2 for No. Working years in the last job and total are also considered as the greatest integer returns.

The codes for environment and place of the accidents and workplaces, and the cause, type, and place in the body of injuries are predetermined by SGK. The variation intervals of the parameters are given in Table 2. Since the variation intervals of the parameters are large, we consider the logarithmic return of each parameter to determine the respected components.

Table 2. *Variation interval of the parameters predetermined by Social Security Institution of Republic of Turkey*

Parameter	Interval
Environment of the Accident	110–220
Place of the Accident	100–200
Workplace	1–9
Environment of the Workplace	11–129
Cause of the Injury	11–89
Type of the Injury	11–120
Place of Injury in the Body	11–78

3.2. The Network and Minimum Spanning Trees

The subject data we used involves randomly chosen 1056 registered occupational accidents in Turkey in between 2013 and 2014 with the aforementioned economic activity group restriction. The correlation distance matrix of each vector standing for occupational accidents is given in Figure 2.

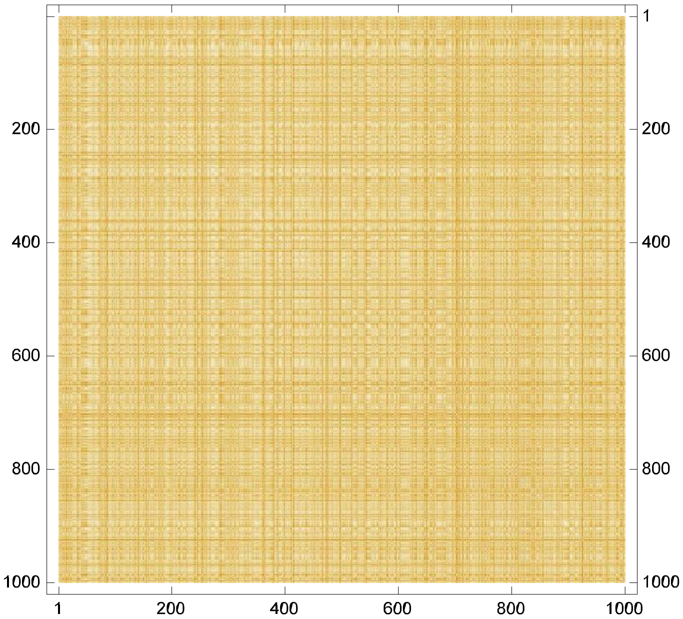

Figure 2. *The correlation distance matrix of randomly chosen data.*

To determine the threshold value, we set $h = 10000$ and resulted TV is 0.00275. This TV value also consistent with our hypothesis that is the lower dimensional embedded occupational accident data involves higher correlations. The network of 1056 agents interacting with $TV = 0.00275$ involves 4 communities as presented in Figure 3. It can also be seen in this figure that each community involves dense connections internally and the Community 1 and Community 2 have a strong connection between them whilst the Communities 2 and 3 are strongly connected.

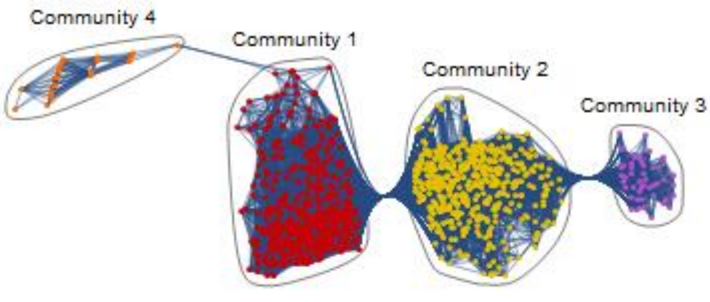

Figure 3. *Communities of the network*

Partition process of the data into sub-communities is given in Figure 4. The process is iterated to 4th level. The highest node in the partition tree is the data and the lowest nodes are the sub-communities in 4th level. Each leaf of the trees is the sub-communities where partition process ends. In Table 3, number of nodes in each sub-community is presented.

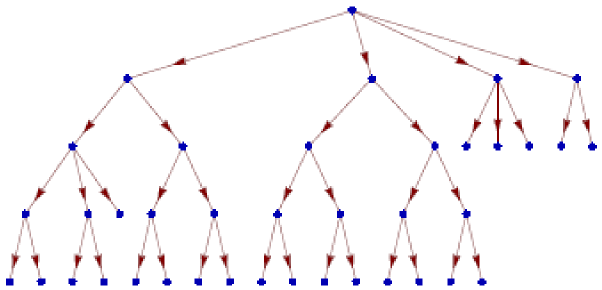

Figure 4. *Sub-communities partition of the data*

Table 3. *The numbers of nodes (N) in each sub-community*

Level	**Community**	**Vertex Number**	**Level**	**Community**	**Vertex Number**
1	1	489	**1**	3	52
	2	485		4	30
2	1.1	282	**3**	1.1.1	151
	1.2	185		1.1.2	129
	2.1	248		1.1.3	2
	2.2	215		1.2.1	98
	3.1	32		1.2.2	87
	3.2	29		2.1.1	125
	3.3	17		2.1.2	123
	4.1	28		2.2.1	119
	4.2	20		2.2.2	96
4	1.1.1.1	78	**4**	2.1.1.1	67
	1.1.1.2	73		2.1.1.2	58
	1.1.2.1	66		2.1.2.1	64
	1.1.2.2	63		2.1.2.2	59
	1.2.1.1	53		2.2.1.1	64
	1.2.1.2	45		2.2.1.2	55
	1.2.2.1	45		2.2.2.1	52
	1.2.2.2	42		2.2.2.2	44

Now, to obtain hierarchies in each community, we first consider the related distance matrix where nodes are the vectors representing the respected occupational accident. The links are formed with the aforementioned formation rule with $TV = 0.00275$ and weights are determined by the d_{corr} similar to the general network construction. Then, we obtain weighted minimum spanning trees in each community by using Kruskal Algorithm [10]. The communities we consider are the leaves of the partition tree presented in Figure 4. The resulted trees are given in Figure 5–7. In Figure 5, the upper row is the MSTs of the sub-communities 3.1, 3.2, 3.3 while the lower row is MSTs of sub-communities 4.1, 4.2, respectively.

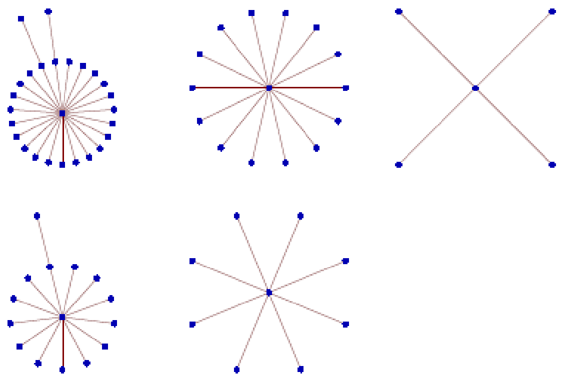

Figure 5. *MSTs of the communities in the leaves of 2nd level of the partition*

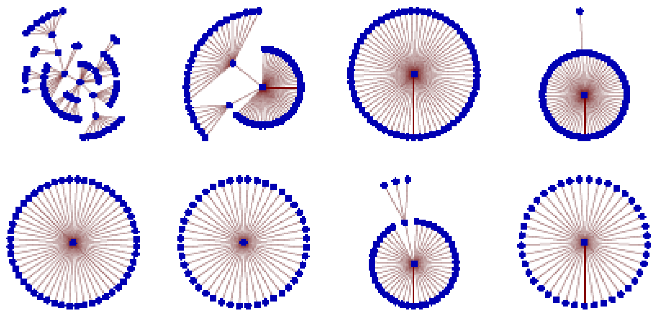

Figure 6. *MSTs of the communities in the leaves of 4th level of the partition rooted to Community 1*

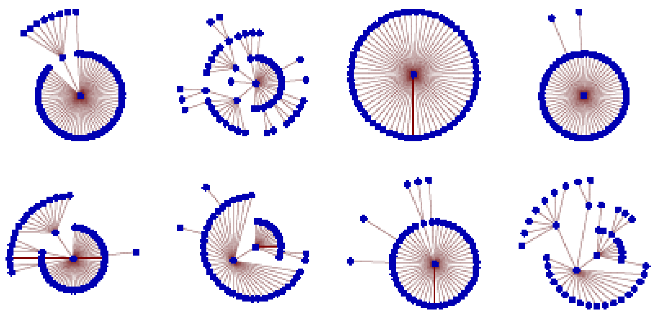

Figure 7. *MSTs of the communities in the leaves of 4th level of the partition rooted to Community 2*

It is straight forward from the figures that, strongly correlated nodes are adjacent to a central node or nodes in each hierarchy. Hence, each central node can be considered as the dominant node in the hierarchy. These dominant nodes are the occupational accidents which is a junction in the information flow. In Tables 4–6 we present central nodes as vectors for each sub-community.

Table 4. *Central nodes in MSTs of Level 2*

Comm	No	Central Nodes in MST
3.1	1	(22, 1, 1, 1, 3, 4, 2.08, 2, 0, 1.0414, 1.7075, 1.3222, 1.7853)
3.2	2	(45, 2, 1, 1, 5, 5, 2.08, 2, 0, 1.0414, 1.6902, 1.0414, 1.7993)
3.3	3	(43, 1, 1, 1, 4, 7, 2.08, 2, 0, 1.0414, 1.7993, 1.3424, 1.716)
4.1	4	(44, 1, 1, 1, 29, 29, 2.08, 2, 0, 1.0414, 1.8512, 1.5051, 1.7993)
	5	(51, 1, 1, 1, 35, 35, 2.08, 2, 0, 1.0414, 1.7924, 1.0414, 1.7324)
4.2	6	(37, 1, 1, 1, 20, 20, 2.08, 2, 0, 1.0414, 1.7853, 1.0414, 1.7324)

Table 5. *Central nodes in MSTs of Level 4 rooted to Community 1*

Comm	No	Central Nodes in MSTs
1.1.1.1	1	(44, 1, 1, 1, 17, 26, 2.08, 2, 0, 1.0414, 1.0791, 1.8388, 1.7242)
	2	(36, 1, 1, 1, 13, 20, 2.08, 2, 0, 1.0414, 1.7075, 1.0791, 1.8388)
1.1.1.2	3	(36, 1, 1, 1, 8, 17, 2.08, 2, 0, 2.0086, 1.6232, 1.0791, 1.7324)
1.1.2.1	4	(44, 1, 2, 1, 4, 21, 2.08, 2, 0, 1.0414, 1.4913, 1.0414, 1.7406)
1.1.2.2	5	(50, 1, 1, 1, 6, 26, 2.08, 2, 0, 1.0414, 1.4913, 1.5051, 1.8129)
	6	(36, 1, 2, 1, 5, 17, 2.08, 2, 0, 1.0414, 1.7075, 1.0414, 1.7242)
1.2.1.1	7	(41, 1, 1, 1, 5, 25, 2.08, 2, 0, 1.1139, 1.6232, 1.5051, 1.8061)
1.2.1.2	8	(36, 2, 1, 1, 3, 20, 2.08, 2, 0, 1.0414, 1.7993, 1.0414, 1.7324)
1.2.2.1	9	(47, 1, 2, 1, 3, 30, 2.08, 2, 0, 1.0791, 1.7924, 1.5051, 1.7324)
	10	(62, 1, 1, 1, 3, 40, 2.08, 2, 0, 1.0414, 1.5051, 1.0414, 1.2787)
1.2.2.2	11	(44, 1, 1, 1, 7, 29, 2.08, 2, 0, 1.0414, 1.591, 1.0414, 1.7242)

Table 6. *Central nodes in MSTs of Level 4 rooted to Community 2*

Comm	No	Central Nodes in MSTs
2.1.1.1	1	(27, 1, 1, 1, 5, 10, 2.08, 2, 0, 1.0414, 1.4913, 1.0414, 1.2552)
2.1.1.2	2	(32, 1, 1, 1, 10, 14, 2.08, 2, 0, 2.0086, 1.6232, 1.5051, 1.8061)
2.1.2.1	3	(52, 2, 1, 1, 5, 21, 2.08, 2, 0, 1.0414, 1.7993, 1.0791, 1.7634)
2.1.2.2	4	(26, 1, 1, 1, 4, 12, 2.08, 2, 0, 1.0414, 1.7924, 1.07918, 1.7242)
2.2.1.1	5	(25, 1, 1, 2, 4, 8, 2.08, 2, 0.3, 1.3617, 1.4913, 1.5051, 1.8129)
2.2.1.2	6	(27, 2, 1, 1, 3, 9, 2.08, 2, 0, 1.0414, 1.4913, 1.5051, 1.7993)
2.2.2.1	7	(35, 1, 1, 1, 4, 10, 2.08, 2, 0, 1.0414, 1.4913, 1.0414, 1.8061)
2.2.2.2	8	(24, 2, 1, 1, 6, 6, 2.08, 2, 0, 1.0414, 1.716, 1.0414, 1.0791)

3.3. Dependence Analysis

Beside the general interpretation of the central nodes in each sub-community, it is also possible to analyse dependence of each node. For this purpose, we perform independence tests among vectors given in Tables 4–6 that are not assuming the normality for the vectors. The tests we use are Blomqvist β, Goodman-Kruskal γ, Hoeffding D, and Kendall τ [14]. A small p-value of each test suggests that it is unlikely that vectors are independent.

In Figures 8–10, the p-values of central nodes of MSTs in Level 2 and Level 4 respect to corresponding dependence tests are presented.

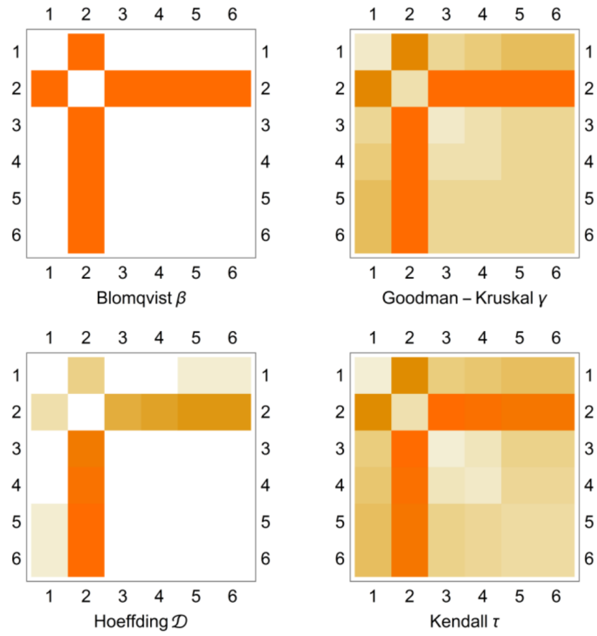

Figure 8. *p-values of each test for central nodes in the sub-communities of Level 2*

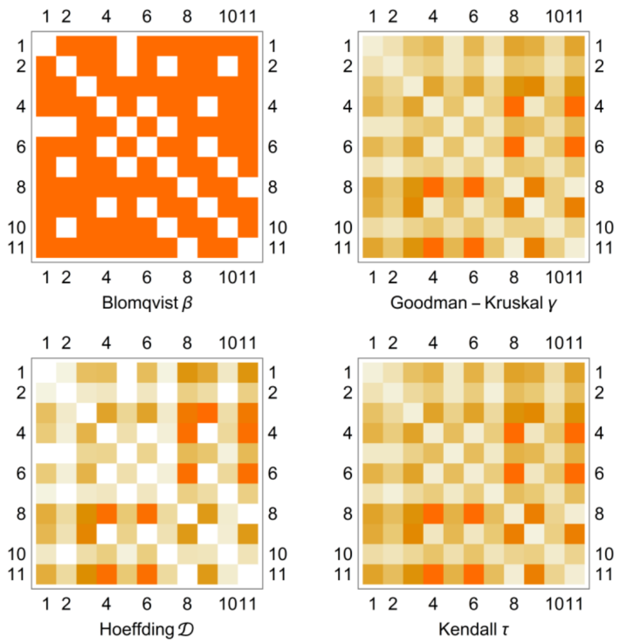

Figure 9. *p-values of each tests for central nodes in the sub-communities of Level 4 rooted to Community 1*

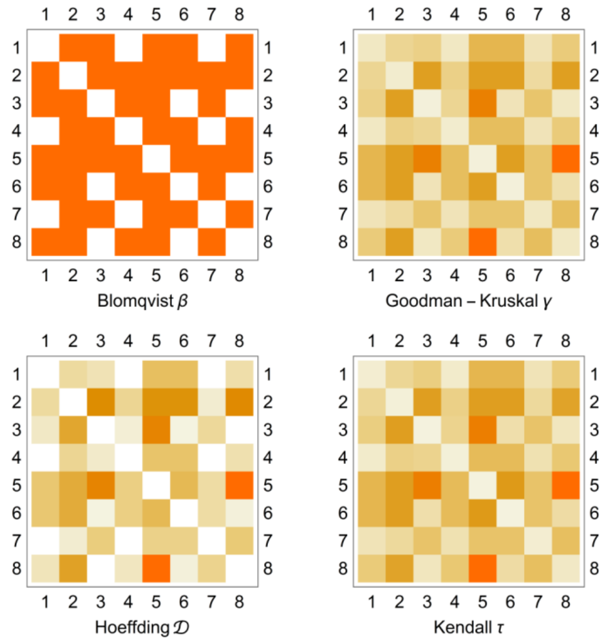

Figure 10. *p-values of each tests for central nodes in the sub-communities of Level 4 rooted to Community 2*

3.4. Topology Analysis

In this section, we present topological measures to analyse MSTs of each sub-community. The mean correlation measure based on $N \times N$ correlation distance matrix $D = [d_{ij}]$ is defined as

$$L_{MCM} = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} d_{ij},$$

where N is the number of nodes in MST.

Another measure to study the property in the MST is normalized tree length is defined as

$$L_{NTL} = \frac{1}{N-1} \sum_{d_{ij} \in \Omega} d_{ij},$$

where Ω is the set of edges, and $N - 1$ denotes the number of edges present in the MST [22,35].

Characteristic path length is used to quantify the average minimal route between pairs of nodes. For an unweighted MST it is defined by

$$L_{CPL} = \frac{1}{N(N-1)} \sum_{i,j:i \neq j} l_{ij},$$

where l_{ij} is the number of edges in the shortest path between nodes i and j [47].

The mean occupation layer is the measurement of the change in the density of the MST. With the central node v_c whose level is taken as zero, the mean occupation layer is defined as

$$L_{MOL} = \frac{1}{N}\sum_{i=1}^{N} lev(v_i),$$

where $lev(v_i)$ denotes the level of node v_i with respect to v_c [34].

The numerical results of the indicated measures of each MSTs are given in Tables 7–9. For the MSTs having more than one central node L_{MOL} is calculated as the mean of each center.

Table 7. *Topological measures of MSTs in Level 2*

Comm	L_{MCM}	L_{NLT}	L_{CPL}	L_{MOL}
3.1	0.00184908	0.00278226	3.10462	1.03846
3.2	0.00140774	0.00286405	2.9381	0.933333
3.3	0.00132076	0.00202056	2.85	0.8
4.1	0.000823653	0.000632966	3.05238	1.36667
4.2	0.00134149	0.00189874	2.90278	0.888889

Table 8. *Topological measures of MSTs in Level 4 rooted to Community 1*

Comm	L_{MCM}	L_{NLT}	L_{CPL}	L_{MOL}
1.1.1.1	0.00771155	0.00353839	4.90709	2.17949
1.1.1.2	0.00262521	0.00268449	3.52302	1.34247
1.1.2.1	0.00115069	0.00212264	2.98508	0.984848
1.1.2.2	0.00127908	0.00172653	3.0151	1.46825
1.2.1.1	0.000868485	0.00140332	2.98149	0.981132
1.2.1.2	0.000908229	0.00221835	2.97828	0.977778
1.2.2.1	0.0010936	0.00221297	3.09949	1.45556
1.2.2.2	0.00109943	0.00169312	2.97677	0.97619

Table 9. *Topological measures of MSTs in Level 4 rooted to Community 2*

Comm	L_{MCM}	L_{NLT}	L_{CPL}	L_{MOL}
2.1.1.1	0.00162081	0.00254792	3.19787	1.10448
2.1.1.2	0.00397509	0.00377414	4.12402	1.63793
2.1.2.1	0.0012565	0.00265351	2.98462	0.984375
2.1.2.2	0.00118569	0.00247417	3.0488	1.01695
2.2.1.1	0.00260599	0.00346775	3.47073	1.26563
2.2.1.2	0.00272482	0.00321681	3.533	1.36364
2.2.2.1	0.00218773	0.00270553	3.1644	1.07692
2.2.2.2	0.00277256	0.00381506	3.94609	1.93182

4. DISCUSSION

Occupational accidents are not dependent on a single factor, but on a complex system of several factors. In this study, we consider a social network of occupational accidents represented as a vector whose entries are different parameters such as Age, Gender, Technical Training, Occupational Safety Training, Working Year in the Last Job, Working Year in the Total, Environment of the Accident, Place of the Accident, Workplace, Environment of the Workplace, Cause of the Injury, Type of the Injury, Place of Injury in the Body.

We restrict analyse to the 10 economic activity groups whose incidence rates are above the Turkey's average. Among these activity groups we randomly choose 1000 registered occupational accidents in Turkey in between 2013 and 2014 calendar years.

To determine communities which are the sets of densely connected nodes we first build the undirected simple graph representation of the network. For the fraction size 10000, the criteria for link forming is determined as 0.00275. This threshold value also indicates that the occupational accidents we consider are highly correlated. Hence, rather than considering the communities in first level, we choose to obtain subcommunities in coarser levels. To obtain sub-communities, we iteratively apply the network forming rule to each sub-community. Since we have a fixed threshold value emerge from the first network and lesser number of nodes in each sub-community, this iteration process runs with exponentially decaying time complexity. In the second level of partition, the communities 3 and 4 end up with three and two sub-communities respectively. In the third level, there occurs the sub-community of 1.1 with 2 nodes. This can be considered as an outlier, so we exclude it from our analysis. After the partition process continuous till the fourth level, the communities 1 and 2 each end up with eight sub-communities.

The central nodes of each MSTs are the most influential occupational accidents in corresponding community which is the local clustering. In the set of central nodes of MSTs of 2nd level of the partition it comes out that the means of the age of workers, the working year in the last job, and the total working year are 40.33, 16, 16.66, respectively. Only one worker is female and all workers have the technical and occupational safety training. All accidents have happened while working in the permanent workplace. This permanent workplace is classified as the production area for all accidents. The cause of the injury varies as contact with a sharp metal tool, hit by an object, muscle crush and physical pressure on the musculoskeletal system. The most frequent type of the injury is superficial wounds on fingers and ankles. In the MSTs of 4th level rooted to Community 1, the means of the age of workers, the working year in the last job, and the total working year are 43.27, 6.72, 24.63, respectively. Only one worker is female and all workers have occupational safety training. Three workers do not have any technical training. All accidents have happened while working in the permanent workplaces. These permanent workplaces are classified as the production, maintenance–repair, and logistic loading and unloading areas. The most frequent causes of the accident are vertical crush with something, contact with a sharp metal tool, and muscle crush. The most frequent type of the injury is superficial wounds on hands and fingers.

In the MSTs of 4th level rooted to Community 2, the means of the age of workers, the working year in the last job, and the total working year are 31, 5.125, 11.25, respectively. Three workers are female while five workers are male. All of the workers have technical training and only one worker does not have any occupational safety training. All accidents have happened while working in the permanent workplaces. These permanent workplaces are classified as the production, maintenance–repair, and logistic loading and unloading areas. The most frequent causes of the accident are horizontal/vertical crush with something, impact of a falling object, and contact with a sharp metal tool. The most frequent types of the injury are superficial, open wounds and sprains on hands and feet.

One of the efficient ways to analyse central nodes in MSTs is to determine dependency of each central node. For this purpose, we applied the statistical tests. The results briefly shows that the most depended central node tuples are (3, 4), (5, 6) for the Level 2 MSTs, (1, 5), (2, 5), (2, 7), (2, 10), (4, 6), (6, 9), (7, 10), (8, 11) for the Level 4 MSTs rooted to Community 1, and (1, 4), (1, 7), (3, 6), (3, 8), (4, 7), (6, 8) for the Level 4 MSTs rooted to Community 2.

Dominant information flows among the agents are through the central nodes of the MSTs. However, to cluster or classify these agents may not be sufficiently enough to understand the complex structure of the occupational accidents. Since the topologies of each MST are formed by the links we also present measures for the links. Mean similarity measure and normalized tree lengths are the measures for the vulnerability of the MST. For the MSTs in Level 2, sub-communities 3.1 and 3.2 can be considered more the more vulnerable than others since L_{NTL} values are greater than the TV = 0.00275. Besides, the deviation between L_{MSM} and L_{NTL} is very large. Hence, the slight changes in the parameters to determine vectors would directly change the topology of respected MSTs. Higher values of L_{MOL} gives a finer network structure while L_{CPL} is a measurement of the compactness of a network. Hence, among the MSTs of Level 2, the MST of sub-community 4.1 is best one to represent occupational accidents in the topological viewpoint. By following the same idea, MSTs of sub-communities 1.2.2.1 and 2.1.1.1 are the best ones rooted to Community 1 and 2, respectively.

5. CONCLUSIONS

In epitome, we present an efficient graph clustering method to analyse occupational accidents in Turkey and we use the Pearson correlation to calculate the similarity measure between the accident data. For this purpose, we represent each accident as a vector of a social actor. Besides, we obtain sub-communities which can be seen as the cluster of clusters iteratively. Moreover, we present the sub-dominant ultra-metric structures among occupational accidents using the MST method. After that, we analyse the central nodes in each MST which are the dominant in the information flow. Furthermore, we use the topological measures to determine which MST is the best to represent occupational accidents in sub-communities.

Afterwards the topological analysis, we find that some MSTs in clusters are representative. The occupational accidents that appear as the nodes of MST of the sub-community 4.1 have the means of age, the working year in the last job, and the total working year as 47.46, 31.4, 31.4, respectively. The occupational accidents above these means have happened in permanent workplace and ended up with superficial wounds on hands caused by muscle crush most frequently. Similarly, occupational accidents that appear as the nodes of MST of the sub-community 1.2.2.1 have the means of age, the working year in the last job, and the total working year as 49.46, 3.84, 33.24, respectively. The occupational accidents above these means have happened in permanent workplace and ended up with superficial wounds on hands and fingers caused by the impact of falling or freely moving object most frequently. As the last representative MST of the sub-community 2.1.1.1, the nodes have the means of age, the working year in the last job, and the total working year as 30.61, 5.74, 12.28, respectively. This means, the occupational accidents above have happened in permanent workplace and most frequently ended up with superficial wounds and sprains on hands and foot, caused by the contact with a sharp metal tool and hit by a falling object.

In the viewpoint of hierarchical structures, we find that the technical training or training of occupational safety and the gender are not effective parameters on the flow of information. Besides, the workers in their mid-ages and have significant experiences in their job become dominant in the hierarchical structure. Also, the superficial wounds on hands and fingers are the most dominant types of the accidents.

An important contribution of our work is that we use the social network of occupational accidents instead of descriptive statistical methods. Therefore, we expect our method is more effective to extract the hierarchical structure of the complex network of occupational accidents. Our analysis based on Pearson correlations can also be extended to analyse occupational accidents in more specific economy activity groups.

CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

[1] Agarwal, G., Kempe, D. “Modularity-maximizing graph communities via mathematical programming”. The European Physical Journal B 66 (3), 409–418 (2008).

[2] Balcı, M. A., Akgüller, Ö. "Average Weakly Hyperedge Domination Number and Actor-Network Application", International Journal of Modelling and Optimization, 4 (5), 346-349 (2014).

[3] Balcı, M. A., Atmaca ,S. P., Akgüller., Ö. “Graph Theoretical Approach To Detection of Outliers in Point Clouds”. International Journal of Advances in Science Engineering and Technology, 3 (4), 59-61 (2015).

[4] Balcı, M. A., Akgüller, Ö. “Mathematical Morphology on Soft Sets for Application to Metabolic Networks”. In Advanced Computational Methods for Knowledge Engineering, Springer International Publishing, 209–218 (2015).

[5] Balcı, M. A. 2016. “Hierarchies in communities of Borsa Istanbul Stock Exchange”. Hacettepe Journal of Mathematics and Statistics, In Press, doi:10.15672/hjms.201614520777.

[6] Balcı, M. A., Akgüller, Ö. Soft Vibrational Force on Stock Market Networks. Library Journal, 3, e3050 (2016).

[7] Yen, N. Y., Timothy K. S., Qun, J. "LONET: an interactive search network for intelligent lecture path generation." ACM Transactions on Intelligent Systems and Technology (TIST) 4.2 (2013).

[8] Benavides, F. G., Benach, J., Muntaner, C., Delclos, G. L., Catot, N., Amable, M. “Associations between temporary employment and occupational injury: What are the mechanisms?” Occupational and Environmental Medicine 63 (6), 416–421 (2006).

[9] Nolan, D. P. “Loss Prevention and Safety Control: Terms and Definitions”. CRC Press, (2016).

[10] Bollobas, B., Modern graph theory (Vol. 184). Springer Science & Business Media (2013).

[11] Condat, L. “A direct algorithm for 1-D total variation denoising”. IEEE Signal Processing Letters, 20 (11), 1054–1057 (2013).

[12] Christian, M. S., Bradley, J. C., Wallace, J. C., Burke, M. J. “Workplace safety: A meta-analysis of the roles of person and situation factors”. Journal of Applied Psychology 94 (5), 1103–1127 (2009).

[13] Clauset, A., Cristopher, M., Mark, E. J. N. “Hierarchical structure and the prediction of missing links in networks”. Nature 453 (7191), 98–101 (2008).

[14] Durante, F., Sempi, C. “Copula theory: an introduction. In Copula theory and its applications”, Springer Berlin Heidelberg, 3–31 (2010).

[15] Joushin, M. K., Saghafipour, A., Noroozi, M., Soori, H., Morasae, E. K. “Epidemiology of accidents and traumas in qom province in 2010”, Archives of trauma research, 2 (3), 113 (2013).

[16] Green, D. G., Bossomaier, E. J. “Complex systems”, Cambridge University Press (2000).

[17] Hughes, B. P., Newstead, S., Anund, A., Shu, C. C., Falkmer, T. “A review of models relevant to road safety”, Accident Analysis & Prevention, 74, 250-270 (2015).

[18] Esmaeilzadeh, S., Ozcan, E., Capan, N. “Effects of ergonomic intervention on work-related upper extremity musculoskeletal disorders among computer workers: a randomized controlled trial”. International archives of occupational and environmental health, 87 (1), 73-83 (2014).

[19] Dwyer, T. “Life and death at work: industrial accidents as a case of socially produced error”. Springer Science & Business Media (2013).

[20] van der Gaast, J. P., Rietveld, C. A., Gabor, A. F., & Zhang, Y. “A Tabu Search Algorithm for application placement in computer clustering”. Computers & Operations Research, 50, 38-46 (2014).

[21] Jackson, M. O. “Social and economic networks”. Vol. 3. Princeton: Princeton University Press (2008).

[22] Jang, W., Lee, J., Chang, W. “Currency crises and the evolution of foreign exchange market: evidence from minimum spanning tree”, Physica A 390, 707–718 (2011).

[23] Leger, D., Massuel, M. A., Metlaine, A. “Professional correlates of insomnia”. Sleep 29 (2), 171–178 (2006).

[24] Leigh, J. P., Marcin, J. P., Miller, T. R. An Estimate of the U.S. Government’s undercount of nonfatal occupational injuries. Journal of Occupational and Environmental Medicine 46 (1), 10–18 (2004).

[25] Mosler, K. “Depth statistics”. In Robustness and complex data structures (pp. 17-34). Springer Berlin Heidelberg (2013).

[26] Lu, L., Tao, Z. “Link prediction in complex networks: A survey”. Physica A: Statistical Mechanics and its Applications 390 (6), 1150–1170 (2011).

[27] Mares, M. “The saga of minimum spanning trees”. Computer Science Review 2(3), 165–221 (2008).

[28] Mantegna, R. N., “Hierarchical structure in financial markets”. The European Physical Journal B-Condensed Matter and Complex Systems, 11(1), 193–197 (1999).

[29] Zimmer, F. M., Schmidt, M., & Magalhaes, S. G. “Correlated cluster mean-field theory for spin-glass systems”. Physical Review E, 89 (6), 062117 (2014).

[30] Scott, J. “Social network analysis”. Sage (2017).

[31] Juárez, F. Review of the principles of complexity in business administration and application in financial statements. African Journal of Business Management, 8 (2), 48 (2014).

[32] Nelson, D. I., Concha–Barrientos, M., Driscoll, T., Steenland, K., Fingerhut, M., Punnett, L., Corvalan, C., “The global burden of selected occupational diseases and injury risks: Methodology and summary”. American journal of industrial medicine 48(6), 400–418 (2005).

[33] Reason, J. Managing the risks of organizational accidents. Routledge (2016).

[34] Onnela, J.-P., Chakraborti, A., Kaski, K. “Dynamic asset trees and portfolio analysis”, The European Physical Journal B 30, 285–288 (2002).

[35] Onnela, J.-P., Chakraborti, A., Kaski, K. “Dynamics of market correlations: taxonomy and portfolio analysis”, Physical Review E 68, 056110 (2003).

[36] Ozkanlar, A., Clark, A. E. “ChemNetworks: A complex network analysis tool for chemical systems”, Journal of Computational Chemistry 35 (6), 495–505 (2014).

[37] Reese, C. D. “ Occupational health and safety management: a practical approach”. CRC press (2015).

[38] Mai, J. E. “Looking for information: A survey of research on information seeking, needs, and behaviour”. Emerald Group Publishing (2015).

[39] Seo, D. C. “An explicative model of unsafe work behavior”. Safety Science 43 (3), 187–211 (2005).

[40] Strogatz, S. H. “Exploring complex networks”. Nature, 410(6825), 268– 276 (2001).

[41] Stam, C. J., Tewarie, P., Van Dellen, E., Van Straaten, E. C. W., Hillebrand, A., Van Mieghem, P. “The trees and the forest: characterization of complex brain networks with minimum spanning trees”. International Journal of Psychophysiology 92(3), 129–138 (2014).

[42] Rushton, L. “The global burden of occupational disease”. Current Environmental Health Reports, 1-9 (2017).

[43] Van Dam, E. R., Haemers, W. H. “Which graphs are determined by their spectrum?”. Linear Algebra and Its Applications 373, 241–272 (2003).

[44] Wang, G. J., Xie, C., Han, F., Sun, B. “Similarity measure and topology evolution of foreign exchange markets using dynamic time warping method: Evidence from minimal spanning tree”. Physica A: Statistical Mechanics and its Applications, 391(16), 4136–4146 (2012).

[45] Wang, H., Hernandez, J. M., Van Mieghem, P. “Betweenness centrality in a weighted network”. Physical Review E 77(4), 046105 (2008).

[46] Burt, R. S., Kilduff, M., Tasselli, S. Social network analysis: Foundations and frontiers on advantage. Annual review of psychology, 64, 527-547 (2013).

[47] Centola, D., van de Rijt, A. “Choosing your network: Social preferences in an online health community”. Social science & medicine, 125, 19-31 (2015).

[48] Viner, M. D. “Occupational Risk Control: Predicting and Preventing the Unwanted”. Ashgate Publishing, Ltd (2015).