T.C.
YILDIZ TECHNICAL UNIVERSITY
MACHINE FACULTY
INDUSTRIAL ENGINEERING DEPARTMENT

ANALYSING AND SELECTING PROPER TECHNOLOGY FOR AMMUNITION AND BLASTING TECHNOLOGIES BY USING SOCIAL NETWORK ANALYSIS

MESUT YILMAZ
14061036

BSC. THESIS

ADVISER
ASSOC. PROF. SERKAN ALTUNTAŞ

İSTANBUL, 2018

PREFACE

Intellectual property is one of the strategically important resources of firms and countries. Patents, which are the most important asset that enables the measurement of intellectual property, protects the exclusive rights of firms and determines company strategies. In addition to granting exclusive rights, patents are subjected to the process of creating new patents for the sustainability of firms and sectors. The most important activity carried out in this direction is to carry out patent analysis between companies and sectors.

In this study, "ammunition and blasting" sub technologies in defense industry for Turkey, is examined in a way to facilitate the decision-making process and this study suggests alternative suitable technologies to be followed and developed.

In addition, this study is built with completely open sourced applications. Also, written scripts for data preparation is open sourced to conduct analyses for different sectors.

May 2018

Mesut YILMAZ

TABLE of CONTENTS

Page

LIST of ABBREVATIONS

IP	Intellectual Property
IPC	International Patent Classification
MDS	Multidimensional Scaling
SNA	Social Network Analysis
SNM	Social Network Mining
TPE	Turkish Patent Institute
UPC	U.S. Patent Classification
USA	United States of America
WIPO	World Intellectual Property Organization

LIST of FIGURES

LIST of TABLES

Page

ABSTRACT

ANALYSING AND SELECTING PROPER TECHNOLOGY FOR AMMUNITION AND BLASTING TECHNOLOGIES BY USING SOCIAL NETWORK ANALYSIS

Mesut YILMAZ

Department of Industrial Engineering

BSc. Thesis

Adviser: Assoc. Prof. Serkan ALTUNTAŞ

Patents, are assets which have strategic importance that guarantee companies' exclusive rights to use and develop technologies. It has also been seen that patent portfolio of companies, provide practical benefits for companies than just protecting exclusive rights. Over time, the patent portfolio had become one of the factors, affecting company valuation. In this regard, it seen that companies form strategies for patent roadmap as well as activities which aligned with their own vision and mission.

For companies, patent strategies are visible strategies, to investigate competitor activities and building insights for sector. It can be predicted that the companies would invest in which technology and develop related products in accordance with selected technology. For companies that want to sustain their competitive advantage, choosing suitable technology and following emerging technology have gained great importance.

Social network analysis is a graph-based method that includes qualitative and quantitative evaluations. Social network analysis allowing qualitative and quantitative evaluation are preferred because they can be easily interpreted from decision makers and provide visualization that combines two types of analysis.

In this study, ammunition and explosive technologies, one of the sub-fields of the defense industry, were analysed by social network analysis method, taking into consideration classification data of patents registered in TPE database. Written softwares used in this work for data scraping, data cleaning and data processing, are open sourced for reproducibility of this work and conducting analyses for different categorical technology evaluations.

Keywords: Social Network Analysis, Social Network Mining, Patent Analysis, Defense Industry, TPE, Graph Theory, Network Visualization, Qualitative Patent Analysis, Quantitative Patent Analysis, Visualization Methods, Data Scraping, Data Cleaning, Data Processing

ÖZET

SOSYAL AĞ ANALİZİ İLE MÜHİMMAT VE PATLAYICI TEKNOLOJİLERİNİN İNCELENMESİ VE UYGUN TEKNOLOJİ SEÇİMİ

Mesut YILMAZ

Endüstri Mühendisliği Bölümü

Lisans Bitirme Tezi

Tez Danışmanı: Doç. Dr. Serkan ALTUNTAŞ

Patentler, şirketlerin teknoloji kullanımı ve üretimine ilişkin özel haklarını teminat altına alan, stratejik öneme sahip varlıklardır. Şirketlerin patent portföyünün, şirketlere koruma avantajı sağlaması dışında şirketler için pratik faydalar sağladığı da görülmüştür. Zaman içerisinde patent portföyü, şirket değerlemesini etkileyen unsurlardan biri olmaya başlamıştır. Bu doğrultuda şirketlerin kendi vizyon ve misyon doğrultusunda gerçekleştirdiği faaliyetlerin yanı sıra patent yol haritası stratejileri de oluşturduğu görülmüştür.

Patent stratejileri, firmalar için sektörün ve rakip firmaların faaliyetlerinin incelenebildiği görünebilir stratejilerdir. Firmaların hangi teknolojilere yatırım yapacakları ve seçilen teknoloji doğrultusunda ürün geliştirme yapacağı öngörülebilmektedir. Rekabet avantajını korumak isteyen firmalar için teknolojileri takip edilmesi ve uygun teknolojilerin seçimi büyük önem kazanmıştır.

Sosyal ağ analizi, kalitatif ve kantitatif değerlendirmeleri içeren çizge teorisi temelli bir yöntemdir. Merkezilik ölçüleri ve ağ görselleştirmeleri ile kalitatif ve kantitatif değerlendirmeye olanak sağlayan sosyal ağ analizi, karar vericiler açısından kolay yorumlanabilmesi ve iki tür analizi birleştiren bir görselleştirmeye sahip olması sebebiyle tercih edilmektedir.

Bu çalışmada, savunma sanayisinin alt alanlarından biri olan mühimmat ve patlayıcı teknolojileri, TPE veri tabanında kayıtlı olan patentlerin sınıf verileri dikkate alınarak, sosyal ağ analizi yöntemi ile incelenmiştir. Çalışma metodolojisinde kullanılan veri kazıma, veri temizleme ve veri işleme amacıyla yazılan yazılımlar, çalışmanın tekrar edilebilirliği ve farklı teknoloji gruplarının analiz edilebilmesi için açık kaynak olarak sunulmuştur.

Anahtar Kelimeler: Sosyal Ağ Analizi, Sosyal Ağ Madenciliği, Patent Analizi, Savunma Sanayii, TPE, Graf Teorisi, Ağ Görselleştirme, Kalitatif Patent Analizi, Kantitatif Patent Analizi, Görselleştirme Yöntemleri, Veri Kazıma, Veri Temizleme, Veri İşleme

CHAPTER 1

INTRODUCTION

1.1 The Goal of the Project

The goal of the project is to analyze inter-technology relations of ammunition and blasting technologies (F42 class) and to recommend suitable sub-technology alternatives for F42 class technologies, by applying Social Network Analysis over IPC (International Patent Classification) category code relations at patent registrations, which is registered at Turkish Patent Institute, Turkey.

1.2 Technology Management and Patent Analysis

Technology management is an important factor for firms by the aspect of competitive landscape. By the help of intellectual property registrations firms gain exclusive rights on maintaining technologies.

The exclusive rights offered by a registered patent mean that patent management is an important issue within the Management of Technology (MOT) field. Companies need to avoid patent infringements when planning their R&D, or they could face possible patent lawsuits or be required to pay a legal cost for the infringement. [1]

1.3 Literature Review

1.3.1 Patent Analysis Methods

Patent analysis methods help decision makers to plan strategic movements. Qualitative and quantitative analyses, provide different kinds of advantaged information for decision makers.

Patent analysis methods may be varying on both or combining qualitative and quantitative analyses. Investigated methods are; selection of core patent, text mining-based methods [19] like subject-action-object models [2] and kinds of NLP based techniques, technology and patent document clustering, quantity-based analysis, Time-Based Analysis and Trend Maps, Ranking Analysis, Citation Statistical Analysis, patent landscape visualization, technology development map and Patent Map [20] Analysis.

1.3.1.1 Citation Methods

Patent citations are defined as the number of reference to prior patents (backward patent citation), or the citations received from subsequent patents (forward patent citation). Patent citations have been often interpreted as a measure of the knowledge diffusing outward from the patents. [3]

Patent citation methods is very useful for selecting state of art and authority patents. But citation analysis is more adequate for building a patent roadmap especially for company domains as (Jeong & Yoon, 2014 [20]) applied.

1.3.1.2 Relation Methods

Relation methods reveals relations between classification codes of patents for different domains like company, country or IPC class domain.[1]

---

[1] This work is analyzing relations within F42 class domain.

The International Patent Classification (IPC), established by the Strasbourg Agreement 1971, provides for a hierarchical system of language independent symbols for the classification of patents and utility models according to the different areas of technology to which they pertain. [4]

IPC code constitutes from section, class, subclass, main group and sub group hierarchical classes. [5] An example for F42B 15/10 ipc code is seem at Table 1.1, below.

Table 1.1 A Structure Example of IPC Code

Classification	**Symbol**	**Field Description (Title)**
Section	F	Mechanical Engineering; Lighting; Heating; Weapons; Blasting
Class	F42	Ammunition; Blasting
Subclass	F42B	Explosive Charges, E.G. For Blasting; Fireworks; Ammunition
(Main) Group	F42B 15/00	Self-propelled projectiles or missiles, Guided missiles
(Sub) Group	F42B 15/10	Missiles having a trajectory only in the air

1.3.1.3 Descriptive Methods

Descriptive analyses, which are examples of quantitative analysis [6], make easy to compare technologies and to learn the state of technology country-wide.

There is applied descriptives; patent count analysis [3] and UPC count analysis[1] applied by (Chen & Chen, 2011 [3]), IPC code frequency analysis to determining code relationship threshold, IPC code frequency distribution, patent number analysis with time series applied by (Park, Lee & Jun, 2015 [21]).

[1] Chen&Chen's work is targeting local classification system UPC at United States of America.

1.3.2 Visualization

Suitable technology problem is a qualitative analysis problem. [6] Visualization is a method for combining qualitative and quantitative analyses as a noticeable way.

Visualization also give a significant value for decision makers. Kim, Suh & Park [21] cites from (Westphal & Blaxton, 1998) that visualization methods are known as one of the best data mining ways to understand because graphical display methods often offer superior result compared to other conventional techniques. [1]

Patent landscape map visualization is a kind of quantitative analysis, an example seems at Figure 1.1. [18]

Figure 1.1 A Patent Landscape Map Example from Water Sector [18]

1.3.3 Social Network Analysis

Social Network Analysis is an analysis which combines qualitative and quantitative patent classifier metrics like closeness degree and modularity with visualization ability.

Social network analysis explores the relationship (”ties”, “arcs” or “edges”) between actors (“nodes” or “vertices”). Historically, the methodology was focused on the relationship between humans. However, since the underlying algorithms originate from the field of graph theory and are universally applicable, modelling of technical relationships. [7]

Sternitzke, Bartkowski & Schramm, which examines the level of detail in which IPC codes [7] should be examined in examining IPC codes in Social Network Analysis, conveys the confidence values from analyses performed at different levels of detail, for 240 patents as in Table 1.2. (Underlined ones are at 1% significance level; Bold ones is at 5% significance level)

Table 1.2 IPC Code Details and Detail Level Correlations

	Subclass Detail	Main Group Detail	Subgroup Detail
Subclass Detail	**-**	-	-
Main Group detailed	**0.843**	-	-
Subgroup Detail	**0.793**	**0.842**	-
Centrality Betweenness	0.025	0.006	<u>0.142</u>

As statistical terms, centrality betweenness is correlating with only subgroup level detail significantly, rather than subclass detail level (4-digit IPC Codes) and main group detail level (7-digit IPC Codes).

It seems where dataset is larger than 240 (> 240), subclass detail level shows high correlation (> 0.800) with subgroup detail (Full IPC Codes) and is converging to significant relationship with centrality betweenness. For this work, subclass detail level gives significant result and it’s selected for social network analysis.

Quantitative measurements below in Social Network Analysis and Graph Theory, is formulated at formulation (1) and (2). N is standing for vertices number; i, j and k is standing for nodes; $\mathrm{sdist}(\mathrm{i}, \mathrm{j})$ is represent shortest distance calculation between node i and node j, while $\mathrm{sdist}(\mathrm{i}, \mathrm{k}, \mathrm{j})$ represent shortest distance over k node.

1.3.3.1 Degree (Centrality) Measurement

Degree is a simple centrality measure that counts how many neighbors a node has. If the network is directed, we have two versions of the measure: in-degree is the number of in-coming links, or the number of predecessor nodes; out-degree is the number of out-going links, or the number of successor nodes. [8]

1.3.3.2 Closeness Centrality Measurement

Closeness centrality (or closeness) of a node is a measure of centrality in a network, calculated as the sum of the length of the shortest paths between the node and all other nodes in the graph. Thus, the more central a node is, the closer it is to all other nodes. [9]

$$\frac{n-1}{\sum sdist(i,j)}, i,j = 1,2\ldots n \tag{1}$$

1.3.3.3 Betweenness Centrality Measurement

In graph theory, betweenness centrality is a measure of centrality in a graph based on shortest paths. For every pair of vertices in a connected graph, there exists at least one shortest path between the vertices such that either the number of edges that the path passes through (for unweighted graphs) or the sum of the weights of the edges (for weighted graphs) is minimized. The betweenness centrality for each vertex is the number of these shortest paths that pass through the vertex. [10]

$$\sum\sum\frac{sdist(i,k,j)}{sdist(i,j)}, i,k,j = 1,2,\ldots n \tag{2}$$

1.3.3.4 Eigen Centrality Measurement

In graph theory, eigenvector centrality (also called eigen centrality) is a measure of the influence of a node in a network. It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. [11]

Eigen centrality is powerful as its discrimination power which results analysis with higher variance.

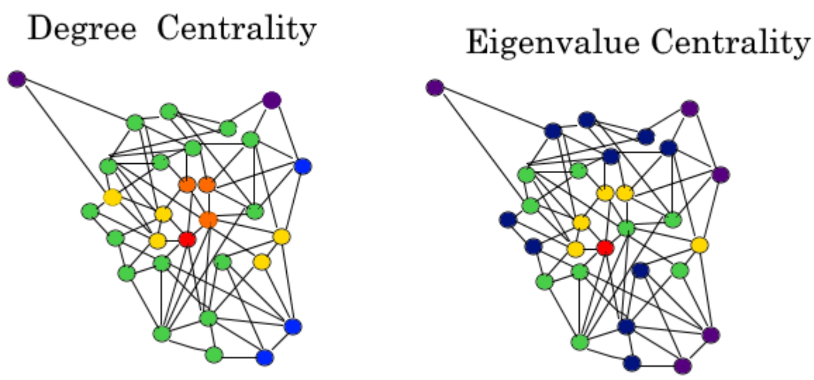

Figure 1.2 Visualized Comparison of Degree Centrality and Eigen Centrality [11]

At Figure 1.2, nodes are colored by their categorized centrality values. Eigen centrality-based network visualization resulted with more colors because of higher variance of eigen centrality. By comparing metrics, degree centrality gives more linear classification categories while eigen centrality gives more logarithmic categories.

1.3.4 Social Network Measurement Correlations

All the measured metrics is expected to correlate among themselves. If metrics doesn't correlate some indications [12] can be foreseen for analyzed network at Table 1.3 below.

Table 1.3 Indications about Discorrelation of Centrality Metrics [12]

	Low Degree	**Low Closeness**	**Low Betweenness**
High Degree	-	Embedded in cluster that is far from the rest of the network	Ego network connections are redundant communication bypasses him/her
High Closeness	Key player tied to important/active alters	-	Probably multiple paths in the network, ego is near many people, but so are many others
High Betweenness	Ego's few ties are crucial for network flow	Ego network monopolizes the ties from a small number of people to many others	-

CHAPTER 2

METHODOLOGY

Technology management and patent analysis methodologies are using end to end approaches from gathering data and pinpointing the result of analysis. Given the problem that social network measures can give different and discorrelated results, it is a problem that which network metrics will be evaluated, and which final technology will be selected. (Park, Lee & Jun, 2015 [21]) recommends the following step-based methodology, taking into account the importance of measures for social network analysis.

Input: Retrieved patent data related to the target technology.

Output: Extracted sustainable technology.

Step 1: Selection of IPC codes

(1.1) Extract all IPC codes from retrieved patent data

(1.2) Select IPC codes with a frequency greater than the threshold value

Step 2: Descriptive statistics of patent data

(2.1) Frequency distribution of patents by IPC codes

(2.2) Yearly trend of the numbers of applied patents

(2.3) Yearly trend of the Numbers of IPC codes included in applied patents

Step 3: Social network mining

(3.1) Visualize technology networking using social network graph

(3.2) Count the degree of top-ranked IPC codes

(3.3) Calculating the closeness centrality of top-ranked IPC codes

(3.4) Calculate the betweenness centrality of top-ranked IPC codes

(3.5) Calculate the graph centrality of top-ranked IPC codes

(3.6) Calculate the shortest distance between top-ranked IPC codes

Step 4: Determine the sustainable technologies based on the results of Step 3

This work constitutes four following steps from data gathering to resulting analysis at Figure 2.1.

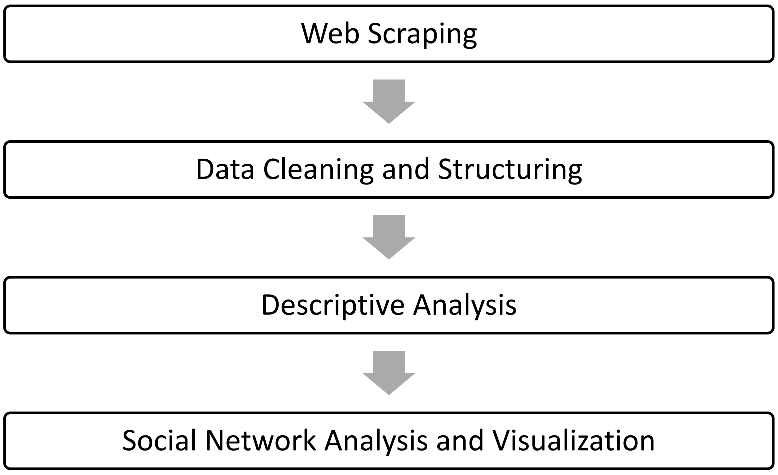

Figure 2.1 Work Methodology

Selection of ipc codes (Step 1) is handled by “Web Scraping” and “Data Cleaning and Structuring” steps. And none of data is handled after a threshold (Step 1.2) Because F42 registrations at TPE is smaller than (Park, Lee & Jun, 2015 [21]) work. Step 2 is processed within “Descriptive Analytics” step and Step 3 is issued under “Social Network Analysis and Visualization” step.

2.1 Web Scraping

Web Scraping (also termed Web Data Extraction, Web Harvesting etc.) is a technique employed to extract large amounts of data from websites whereby the data is extracted [13] and saved to a local file in a computer or to a database or in a table (spreadsheet) format.

Web scrapping algorithms is widely implemented by scripting languages with fully automated processes. In terms of TPE (Turkish Patent Institute) patent search page[1] structure, web scrapping script cannot be run fully automated. TPE patent search page shows three characteristics;

- The page is just a single page, which shows result table and patent detail table at the same table
- The page is authenticated with visual captcha methods which require human entry per patent view request
- Search system doesn't come with different endpoints, which makes harder to track iteration for a fully automated script

Because of the characteristics of search page, a client side [14] and browser-based language is suitable for gathering. As a language selection process, a semi-automated JavaScript[2] code[3], which seem at Figure 2.2 while running the code, is implemented to gather data.

---

[1] http://online.turkpatent.gov.tr/EPATENT/servlet/EPreSearchRequestManager

[2] JavaScript's biggest advantage is the ability to run in a web browser interface and the lack of any environment setup process.

[3] Gather script can be found at https://github.com/imesut/f42/blob/master/collector.js

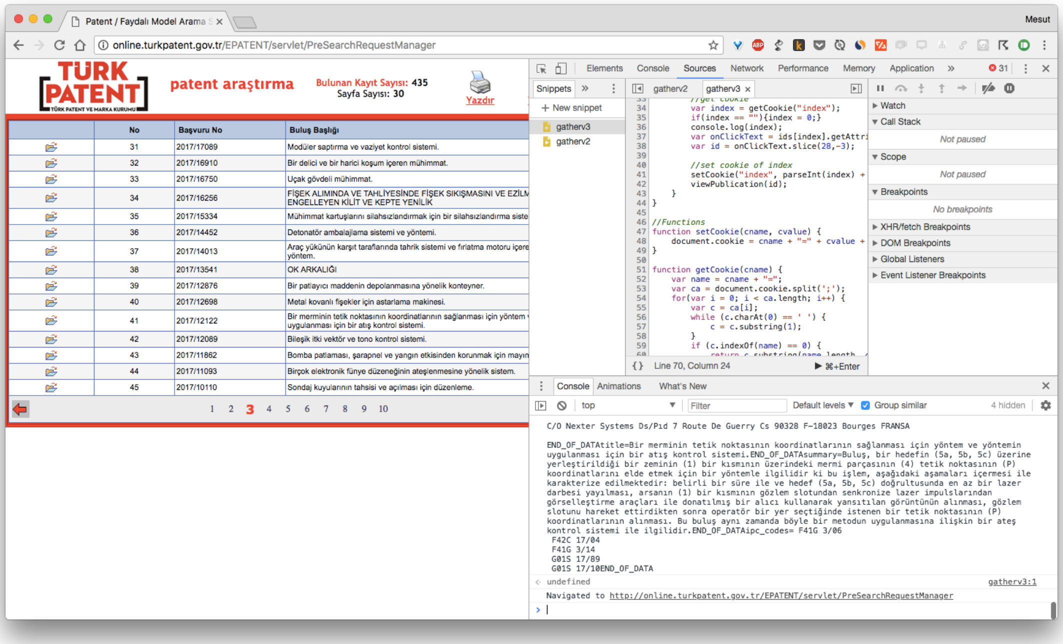

Figure 2.2 Web Scraping Script Collects Information

Because of the output area, JavaScript gathering code can create one dimensional (single line, not separated) data array, below.

---

gatherv3:28 apply_number= Başvuru Numarası: **2018/05433** END_OF_DATA apply_date = Başvuru Tarihi: **2015/08/04** END_OF_DATA registration_id= Tescil Numarası: **2018 05433** END_OF_DATAregistration_date= Tescil Tarihi: **2018/04/24** END_OF_DATA protection_type = Koruma Tipi: **Patent**END_OF_DATAowner_1=**DIEHL DEFENCE GMBH & CO. KG \n Alte Nußdorfer Strasse 13 88662 Überlıngen ALMANYA** \n END_OF_DATAowner_2=**DR. ARNO HAHMA \n Schulstr. 2 De - 91239 Henfenfeld ALMANYA** \n END_OF_DATAtitle=**Aktif maddesi ve (…)**END_OF_DATAsummary=**Buluş, piroteknik olarak bir sis (….)** END_OF_DATA ipc_codes= **F42B 5/15**END_OF_DATA \n gatherv3:1 undefined

---

Figure 2.3 A registration output of web scraping script

At Figure 2.3. required data is showed bold, identifier tags are shown in red. The remaining output, which should be disregarded, is from TPE search page.

2.2 Data Cleaning and Structuring

Browser based web scraping scripts gives one dimensional data. To handle data as structural and easy to process with spreadsheet applications, one dimensional data should be pre-processed for table view.

For this step there is no requirement about running environment. So, Python is one of the suitable language and used for building a cleaner and structuring script.[1][2]

Table 2.1 A Structured Registration Output of Cleaner Script

apply_number	apply_date	registration_id	registration_date	protection_type	…	ipc_codes
2018/05433	2015/08/04	2018 05433	2018/04/24	Patent		F42B 5/15

Also, this script layer process data for calculating reoccurrences and relationship of frequencies -also known as edges table in terms of social network analysis- which will be used at Social Network Analysis step.

Table 2.2 Sample Relationship Data Output from Cleaner Script

Source (IPC)	Target (IPC)	Weight (Count)
F41H	F42B	19
F42D	E21C	14
F42D	G01V	2

2.3 Descriptive Analysis

Descriptive analyses which cited at literature research topic, is processed by Spreadsheet application with Pivot Table features.

---

[1] Will be called as “cleaner” in this work.

[2] Cleaner script can be found at https://github.com/imesut/f42/blob/master/cleaner.py

For refining ownership fields of data, regex (Regular Expression) and character limiting methods are used over data. Regex method is widely used for matching different strings which differ by some characters.

Table 2.3 Different Company Names for Same IP Owner

DİEHL BGT DEFENCE GMBH & CO	DIEHL BGT DEFENCE GMBH & CO
DIEHL DEFENCE GMBH	DİEHL DEFENCE GMBH

Another problem is the diversified ownership which investigated under name normalization [15] methods.

For F42 technology group, building 6-character code names with human control is a convenient way to make data unique about intellectual property owner. After regex and code name generation, the example at Table 2.3 will become like at Table 2.4.

Table 2.4 Normalized Company Names for the Same IP Owner

“diehl ”	“diehl ”
“diehl ”	“diehl ”

2.4 Social Network Analysis and Visualization

At social network analysis step, a visualization software called as Gephi is used for performing social network degree and metrics calculation and for visualizing technology group relations.

Gephi is an open-source software for network visualization and analysis. It helps data analysts to intuitively reveal patterns and trends, highlight outliers and tells stories with their data. It uses a 3D render engine to display large graphs in real-time and to speed up the exploration. [16]

Besides the standard social network measurement, modularity of patent network is analyzed. Modularity (community detection) is a measure of network structure. It was designed to measure the strength of division of a network into modules. Networks with high modularity have dense connections between the nodes within modules but sparse connections between nodes in different modules. [17] Also it is recommended to use "resolution=1.0" and "randomize" parameters in Gephi by (Ji & Machiraju, 2015 [17]).

For visualization, different kind of visualization algorithms is used with different combination. These algorithms and its parameters is shown at Table 2.5.

Table 2.5 Algorithms and Parameters Used in Visualization

Order	**Algorithm**	**Parameters**
1	Fruchterman Reingold	Area = 5000.0 Gravity = 20
2	Force Atlas 2	Tolerance (speed) = 0.1 Scaling = 60.0 Gravity = 2.0 Prevent Overlap = true
3	Noverlap	Speed = 3.0 Ratio = 1.2 Margin = 5.0
4	Label Adjust	-
5	Rotate	Angle = -70

CHAPTER 3

DESCRIPTIVES AND ANALYSES

3.1 Descriptive Analyses

3.1.1 Apply Statistics

There are 391 registered patents, which applied for taking exclusive right of use and taking ownership of IP, related with F42 IPC class of defence industry in TPE database. Table 3.1 shows IP registrations by top 6 applier countries.

Table 3.1 Intellectual Property Registration Numbers by Countries

Country	**Registration Numbers**	**Percentage (1970-2018)**
Germany	91	23,3%
Turkey	81	20,7%
USA	60	15,3%
France	46	11,8%
Swiss	12	3,1%
Spain	11	2,8%

As the trendline in F42 technological development is sustaining to growth, there is a change on ownership percentage of applier countries. Figure 3.1 shows registration numbers of countries by years.

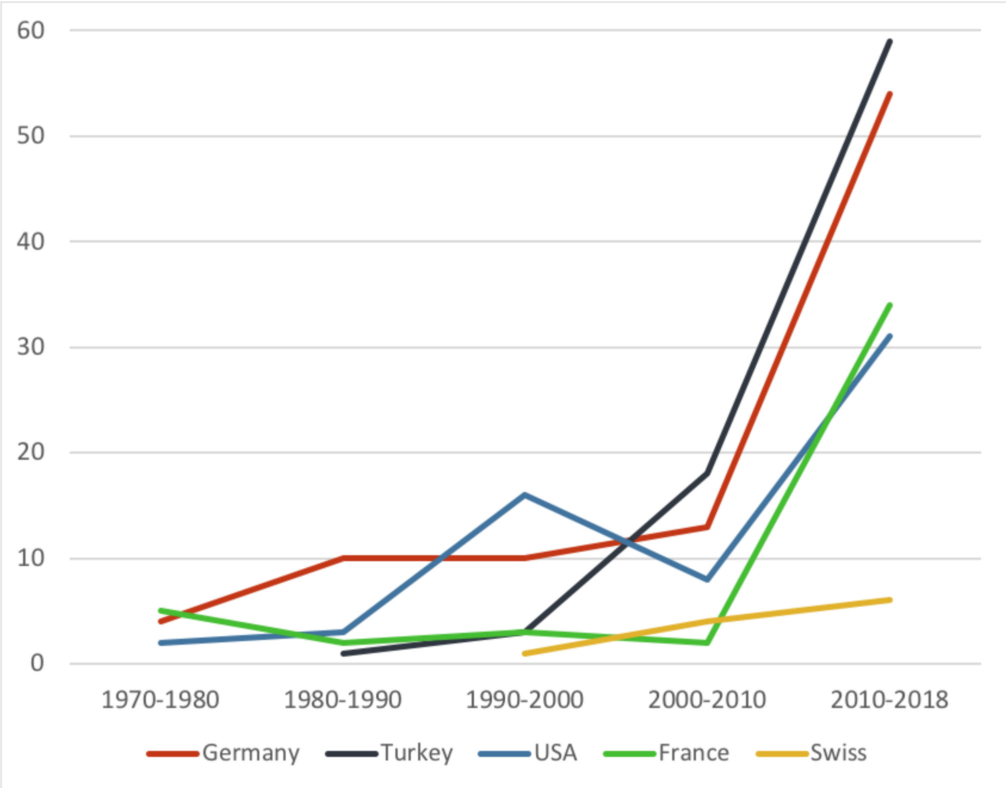

Figure 3.1 Top 5 Country by F42 Class IP Registration number over Years

It is obvious that, defence industry has key importance regarding different strategic aspects. Especially as a precondition for exporting weapons (F41) to a country, exporter firm/country should apply for IP registration in desired country. Related to weapon (F41) technologies, ammunition technologies (F42) also should be protected. After 2000-2010 period Turkey has started to become the biggest applier for F42 technologies. Other biggest appliers are Germany and France-USA followingly.

By strategic perspectives, market dominance is also important for exporters to take exclusive rights on their products. So, exporters and countries can apply policy to take the dominance at market at macro scale. To understand IP owning landscape, the percentage owning by years is stated at Figure 3.2 below.

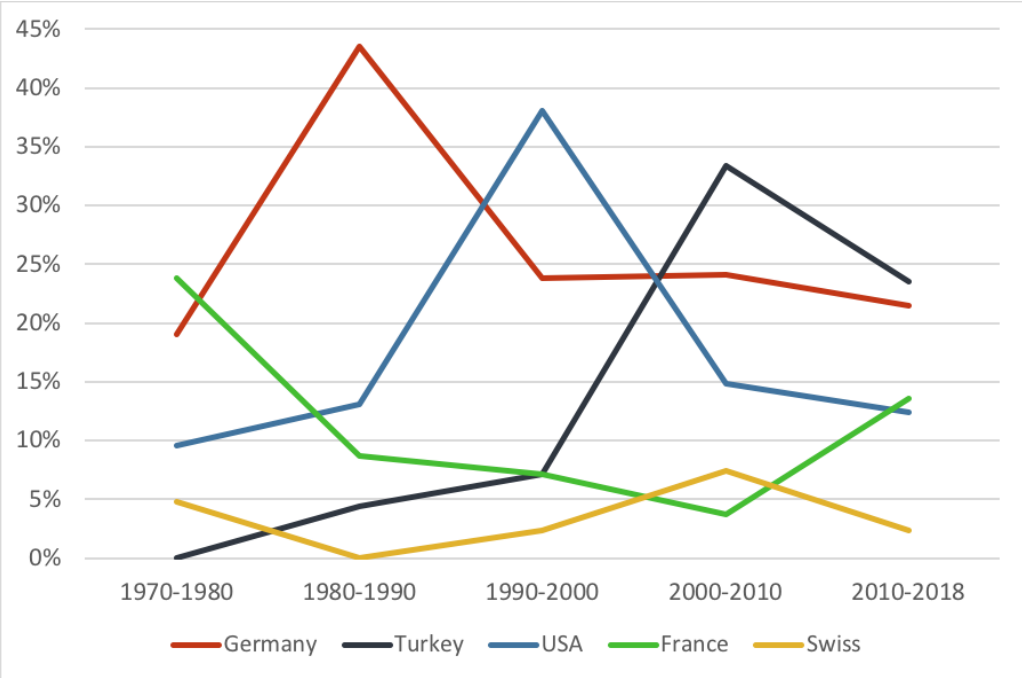

Figure 3.2 Top 5 Country by F42 Class IP Registration Percentage over Years

At lower level, IP owning is a discriminating power for firms. Like countries firms can take a policy for obtaining market share. The biggest IP owner companies' statistics is stated at Table 3.2 and IP owning change by year is stated at Figure 3.3 below.

Table 3.2 IP Registration number by F42 Class Percentage (1970-2018)

Company Name	**Foundation Date**	**Country**	**Percentage**
Rheinmetall	1889	Germany	10,0%
Raytheon Company	1922	USA	5,6%
Nexter Munitions	2006	France	3,3%
Diehl Bgt Defence	1902	Germany	2,3%
Dynamit Nobel Ammotec	1871	Germany	2,3%
Hughes Aircraft Co	1932-1997	USA	2,3%

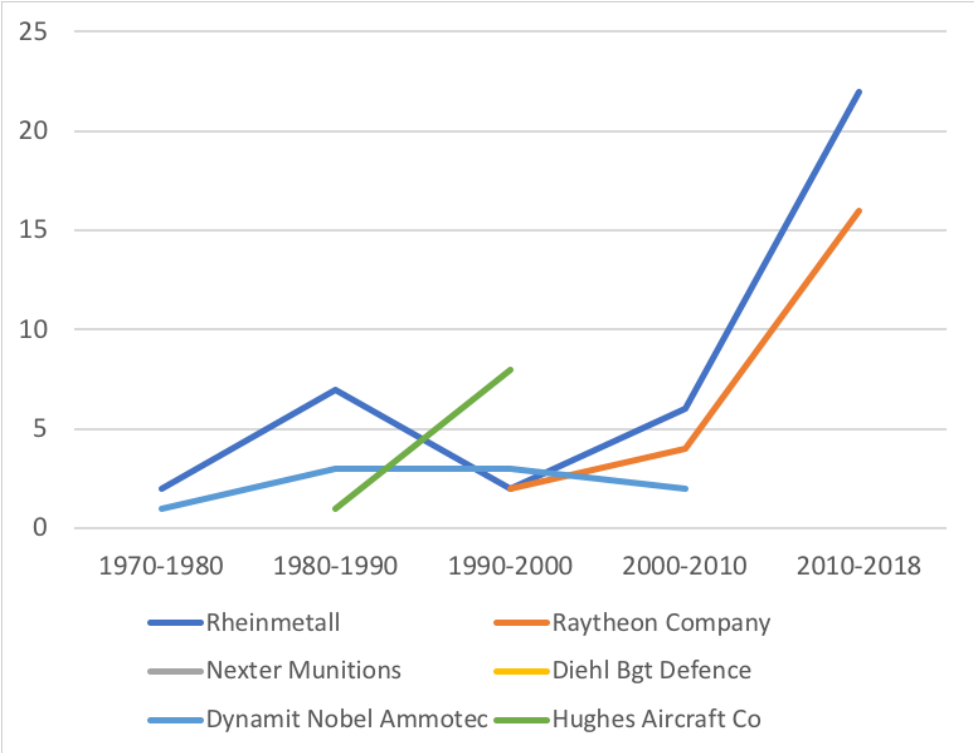

Figure 3.3 Top 6 Company IP Registration Number over Years

It seems that, Rheinmetal, a firm originated at Germany, is a significant actor at F42 technologies in Turkey. The second company, Raytheon, a firm at USA, is a sustainable applier unlikely other following company.

To distinguish know-how level of appliers, (Park, Lee & Jun, 2015 [21]) underlines the importance of total IPC code citation number for companies and countries, from considering the complexity of patent registration where a complex and know-how intense apply consist different categorical fields and intertwined use of different technologies.

Figure 3.4 shows total ipc code citation number by companies. Diehl companies group (Germany) is the leader for ipc code citation, followingly by Rheinmetal (Germany) and Raytheon (USA). It’s possible to find an evidence for Diehl group is applying for more intense technologies more than whole of its portfolio.

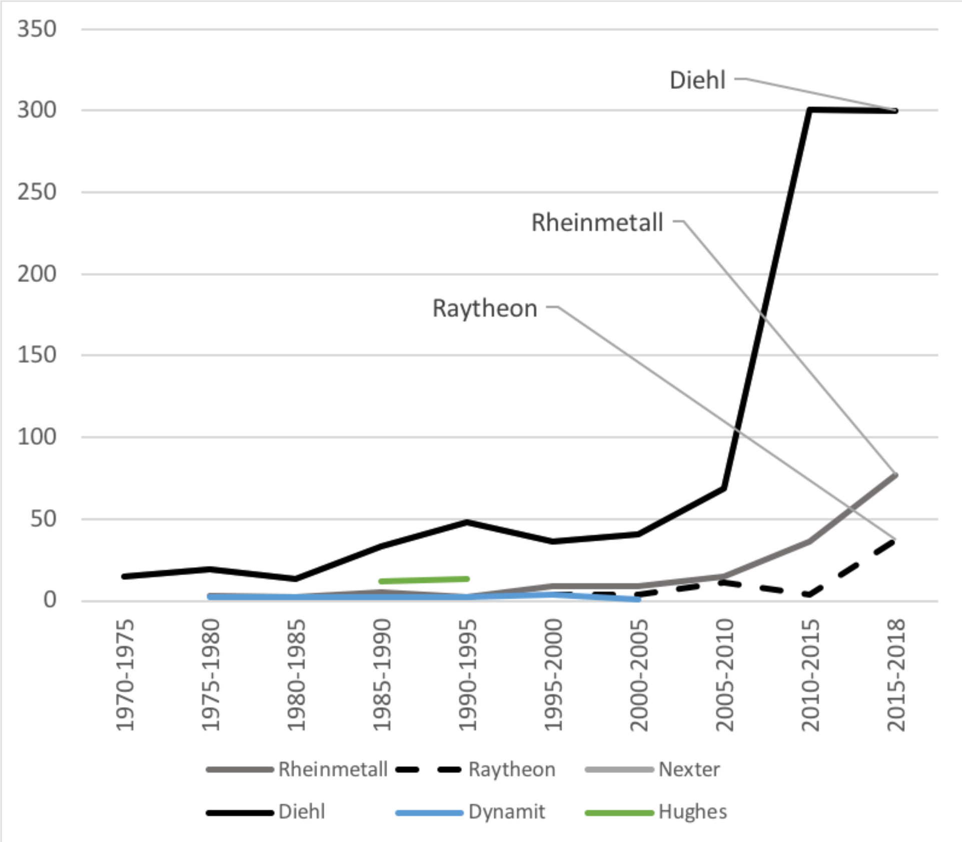

Figure 3.4 Total IPC code referring by top companies over years

3.1.2 Technology Development Statistics

F42 technologies in Turkey, is becoming more complex by the time. Until 2007, F42 technologies has a flat trendline for development (Figure 3.5). After 2007 F42 applies has started to increase as a trendline changer. Especially after 2007, complexity of applies has started to exponentially increase.

As the complexity of patent applications, cited ipc code number is also important. %58 percent (Figure 3.6) of total registration is registered for more than one ipc codes, is comparably lower than a company portfolio. This low level of citation is also trended to increase, after 2010, the ratio of ipc code citation per patent is exceeding, 2 citation level in-average. (Figure 3.7)

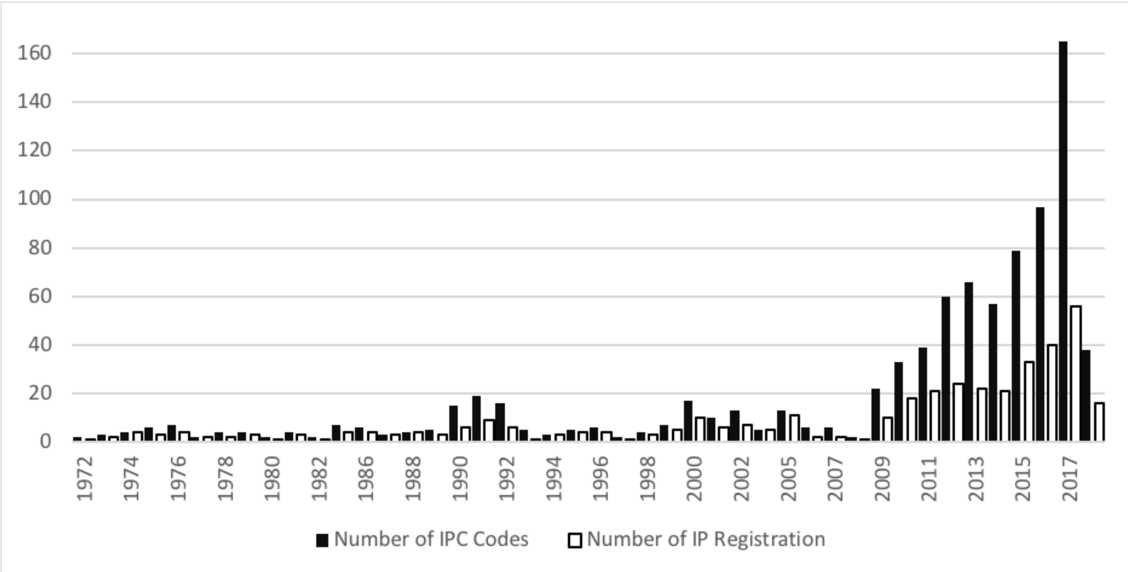

Figure 3.5 Total Number of IPC codes and IP Registrations

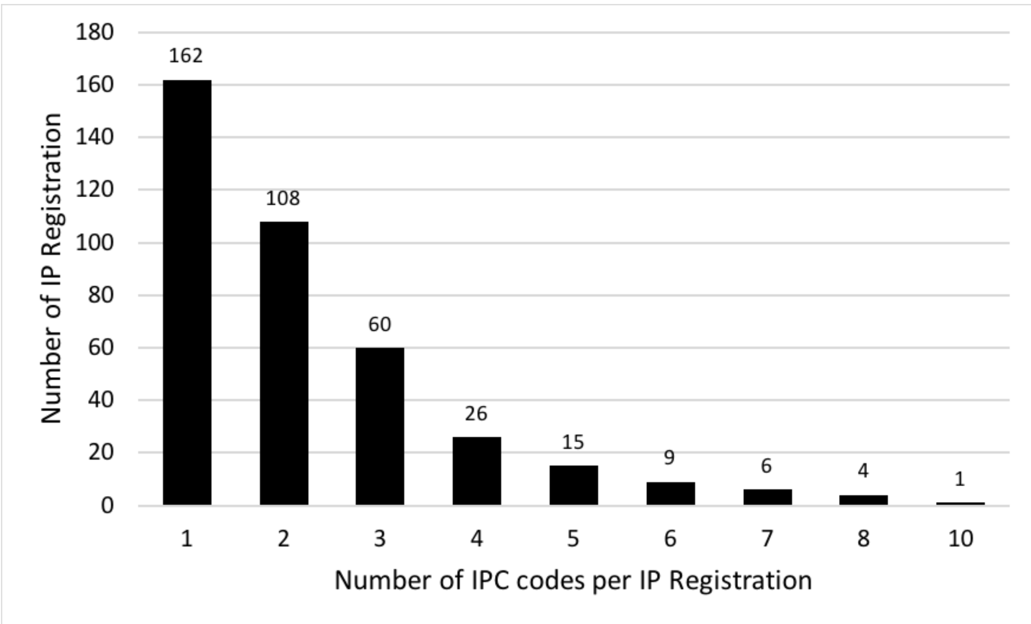

Figure 3.6 Frequency distribution IPC code per patent

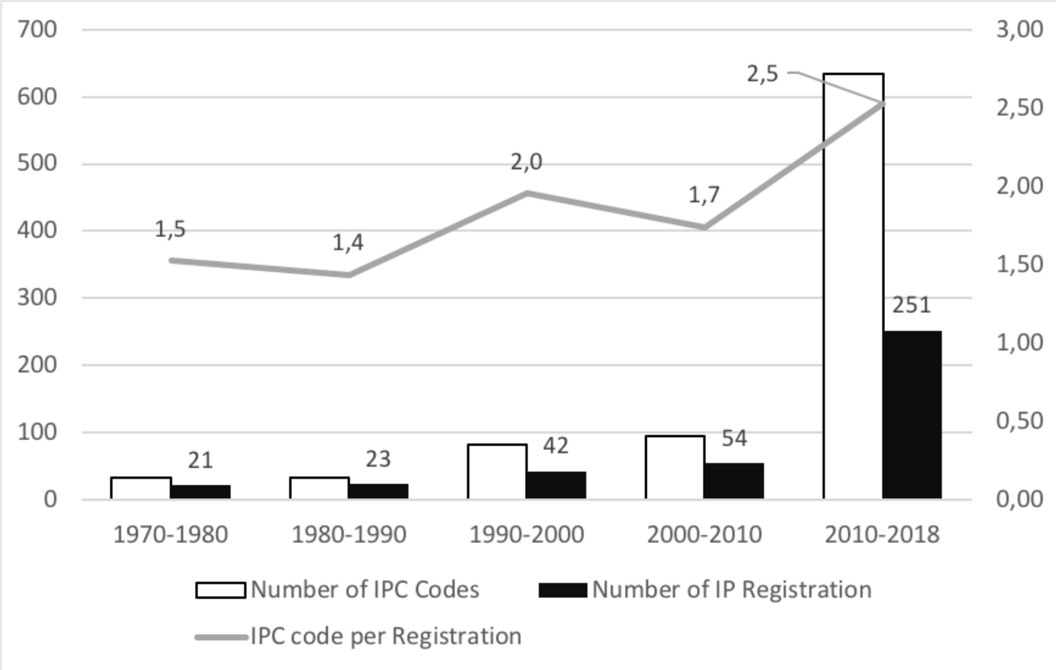

Figure 3.7 Numbers of IP Registrations and IPC codes by year

3.1.3 Sub-Technology Statistics

Related patents with F42, cites different IPC classes as a result of intertwined and more complex applies. The cited class frequencies for registered patents in Turkey is showed at Figure 3.8. Regarding sub class distribution of F42, explosive technologies (F42B) is most demanded sub technology. Ammunition fuze technologies (F42C) and blasting technologies (F42D) is following explosives.

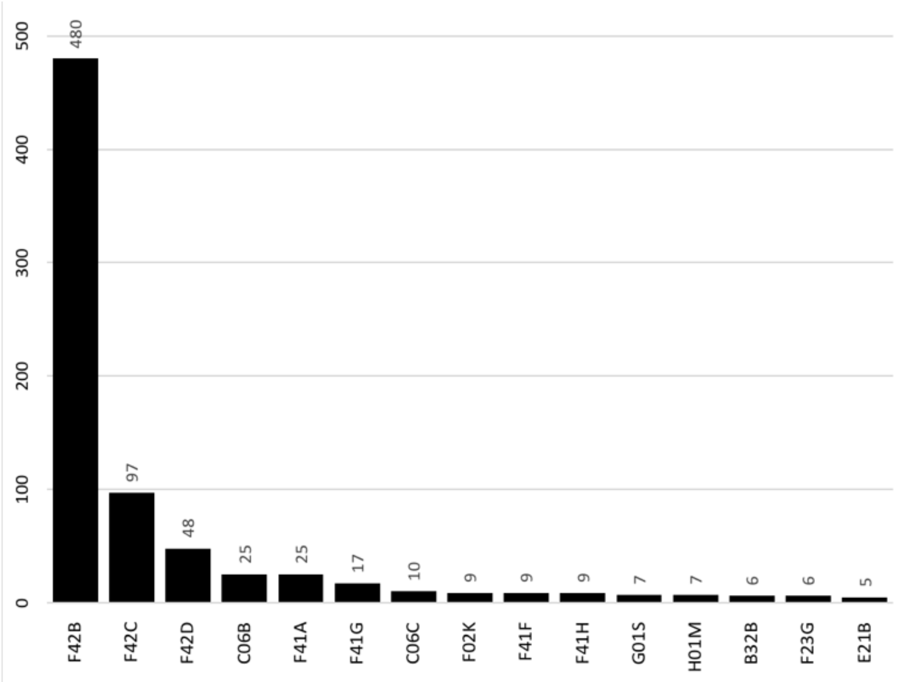

Figure 3.8 IPC codes with frequency

Table 3.3 Top 10 IPC Codes

IPC	Description
F42B	Explosive Charges; Ammunition (Explosive Compositions; Fuzes; Blasting)
F42C	Ammunition Fuzes (Blasting Cartridge Initiators; Chemical Aspects); Arming or Safety Means Therefor (Filling Fuzes; Containers for Fuzes)
F42D	Blasting (Fuses, E.G. Fuse Cords; Blasting Cartridges)
C06B	Explosive or Thermic Compositions; Manufacture Thereof; Use of Single Substances as Explosives
F41A	Functional Features or Details Comton To Both Small Arms and Ordnance, E.G. Cannons; Mountings for Small Arms or Ordnance
F41G	Weapon Sights; Aiming (Optical Aspects)
C06C	Detonating or Priming Devices; Fuses; Chemical Lighters; Pyrophoric Compositions
F02K	Combustion Engines; Hot-Gas or Combustion-Product Engine Plants
F41F	Armour; Armoured Turrets; Armoured Or Armed Vehicles; Means of Attack or Defense, E.G. Camouflage, In General
F41H	Armour; Armoured or Armed Vehicles; Means of Attack or Defence

3.2 Social Network Analysis

Social network analysis is providing two-sided approach. With centrality measures, quantitative analysis can be performed. Also, social network provides powerful visualization. Under this topic, before quantitative analysis which will be at the base of quantitative results, quantitative analyses is performed.

3.2.1 Centrality Measurements (Quantitative Analysis)

For F42 technologies network, centrality measurements are calculated by Gephi's built-in statistical tools, at Table 3.4 below.

Table 3.4 Centrality Measurements

IPC	Degree	Closeness Centrality	Betweenness Centrality	Authority (Hub)	Eigen Centrality
F42B	67	0,868	0,846	0,598	1,000
F41A	15	0,530	0,028	0,222	0,363
F42C	15	0,552	0,099	0,191	0,321
F42D	17	0,560	0,138	0,176	0,307
F41G	10	0,506	0,008	0,147	0,246
F41H	9	0,520	0,026	0,142	0,238
G06K	7	0,488	0,001	0,140	0,226
C06C	7	0,523	0,004	0,133	0,224
B23K	6	0,485	0,000	0,133	0,215
B41M	6	0,485	0,000	0,133	0,215

As (Du, 2016) cites that, a correlation analysis should be performed for centrality measures for network analysis. At Figure 3.9 and Table 3.4, it's seem that, all the correlations can be accepted as highly correlated (> 0.800) referring to inter usability of centrality measures. Betweenness centrality should be analyzed separately because it shows minimum correlation in average with other measures.

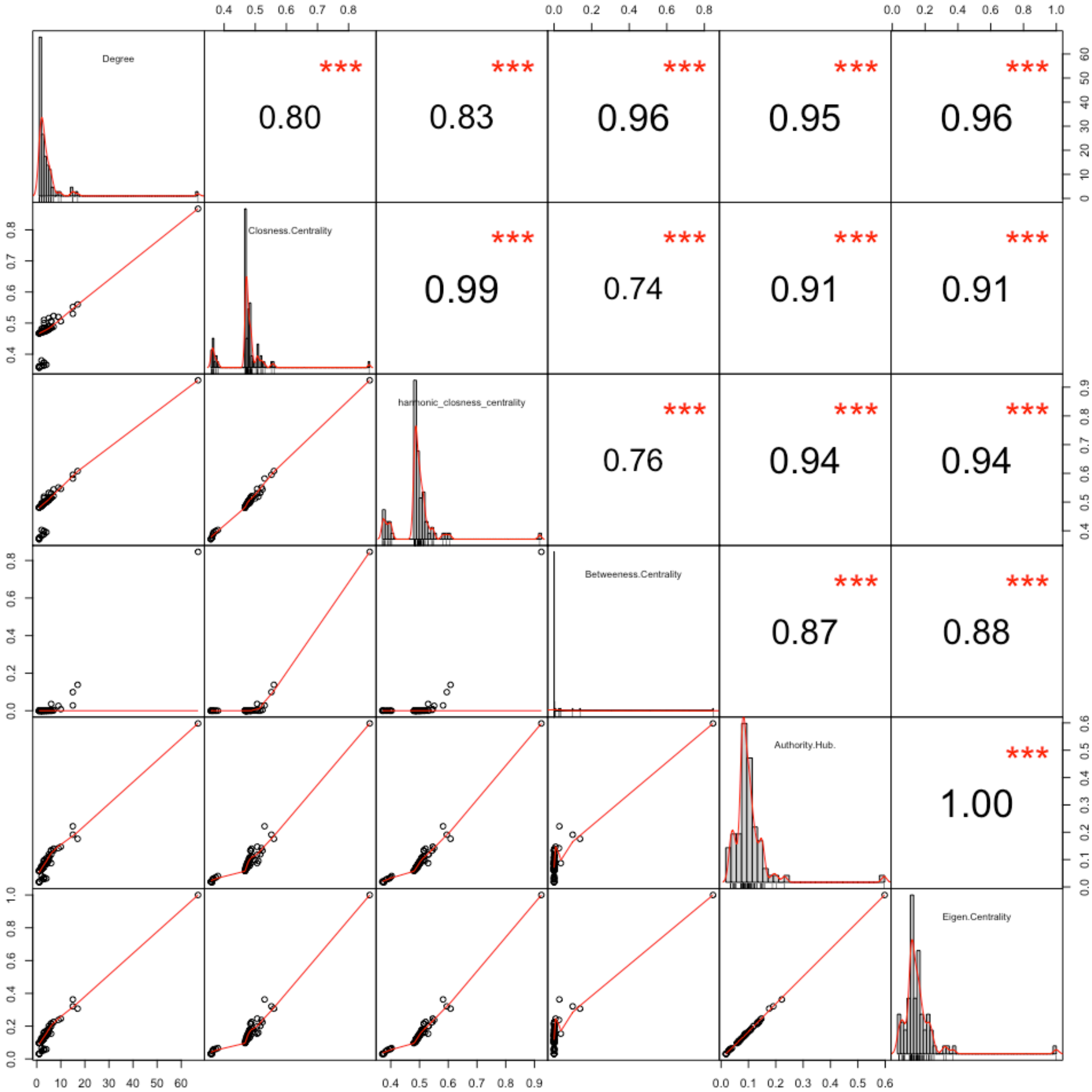

Figure 3.9 Correlation of Centrality Measures

Table 3.5 Correlation Table of Centrality Measures

	Degree	Closeness Centrality	Betweennes s Centrality	Authority (Hub)	Eigen Centrality
Degree	1,000	0,797	0,961	0,953	0,958
Closeness Centrality	0,797	1,000	**0,741**	0,908	0,908
Betweenness Centrality	0,961	0,741	1,000	0,871	0,877
Authority (Hub)	0,953	0,908	0,871	1,000	1,000
Eigen Centrality	0,958	0,908	0,877	1,000	1,000

As the result of correlation analysis, it is possible to take a single parameter for selecting suitable technology. In average, eigen centrality resulting with 0,936 correlation coefficient maxima while betweenness centrality gives minimum coefficient as 0,863. Also, according to (Park, Lee & Jun, 2015 [21]) it is suggested to use step by step approach.

At Table 3.4 data is sorted by eigen centrality value. F42B class has the highest value for all of the measurements. F42B is identified as key technology.

F42D is an outlier example, when degree, closeness and betweenness is higher than F42C class, eigen centrality gives lower value rather than F42C. It can prove that, the relations of F42D may be related with outlier examples also. It is cited at Sub-Technology Statistics section of this document; Sub groups of F42 class is ranked decreasing order like; B-C-D sub groups.

When F41A is compared with F42C, F41A is less close and less between but resulted as more authority and eigen value. It’s means that, F41A is likely to be good portfolio technology with F42 class relation.

When result table is ranked according to degree and eigen centrality, switch between F41A and F42D is occurring. This case is a symptom as, patents are focusing on ammunition rather than blasting. Because, F41A (weapon technology group) can replace a blasting sub group (F42D) easily.

Table 3.6 Top 10 IPC codes by Degree, Rank and Descriptions

Rank	Degree Rank	IPC	Description
1	1	F42B	Explosive Charges, E.G. For Blasting; Fireworks; Ammunition (Explosive Compositions; Fuzes; Blasting)
2	3	F41A	Functional Features or Details Comton To Both Small Arms and Ordnance, E.G. Cannons; Mountings for Small arms Or Ordnance
3	4	F42C	Ammunition Fuzes (Blasting Cartridge Initiators; Chemical Aspects); Arming or Safety Means Therefor (Filling Fuzes; Fitting or Extracting Primers in Or from Fuzes; Containers for Fuzes)
4	2	F42D	Blasting (Fuses, E.G. Fuse Cords; Blasting Cartridges)
5	5	F41G	Weapon Sights; Aiming (Optical Aspects)
6	6	F41H	Armour; Armoured Turrets; Armoured Or Armed Vehicles; Means of Attack or Defense, E.G. Camouflage, In General
7	7	G06K	Recognition of Data; Presentation of Data; Record Carriers; Handling Record Carriers
8	8	C06C	Detonating or Priming Devices; Fuses; Chemical Lighters; Pyrophoric Compositions
9	9	B23K	Soldering or Unsoldering; Welding; Cladding or Plating by Soldering or Welding; Cutting by Applying Heat Locally; Working by Laser Beam
10	10	B41M	Printing, Duplicating, Marking, Or Copying Processes; Color Printing (Correction of Typographical Errors; Processes for Applying Transfer Pictures; Fluid Media for Correction of Typographical Errors by Coating; Printing Textiles)

3.2.2 Graph Analysis (Qualitative Analysis)

At following pages there are different kinds of visualization examples which focus on different measurements. Each of visualization is described with centrality measures also by the aspect of technology selecting problem.

Graphs is colored by modularity classification algorithm. Same node color shows the same group for different nodes.

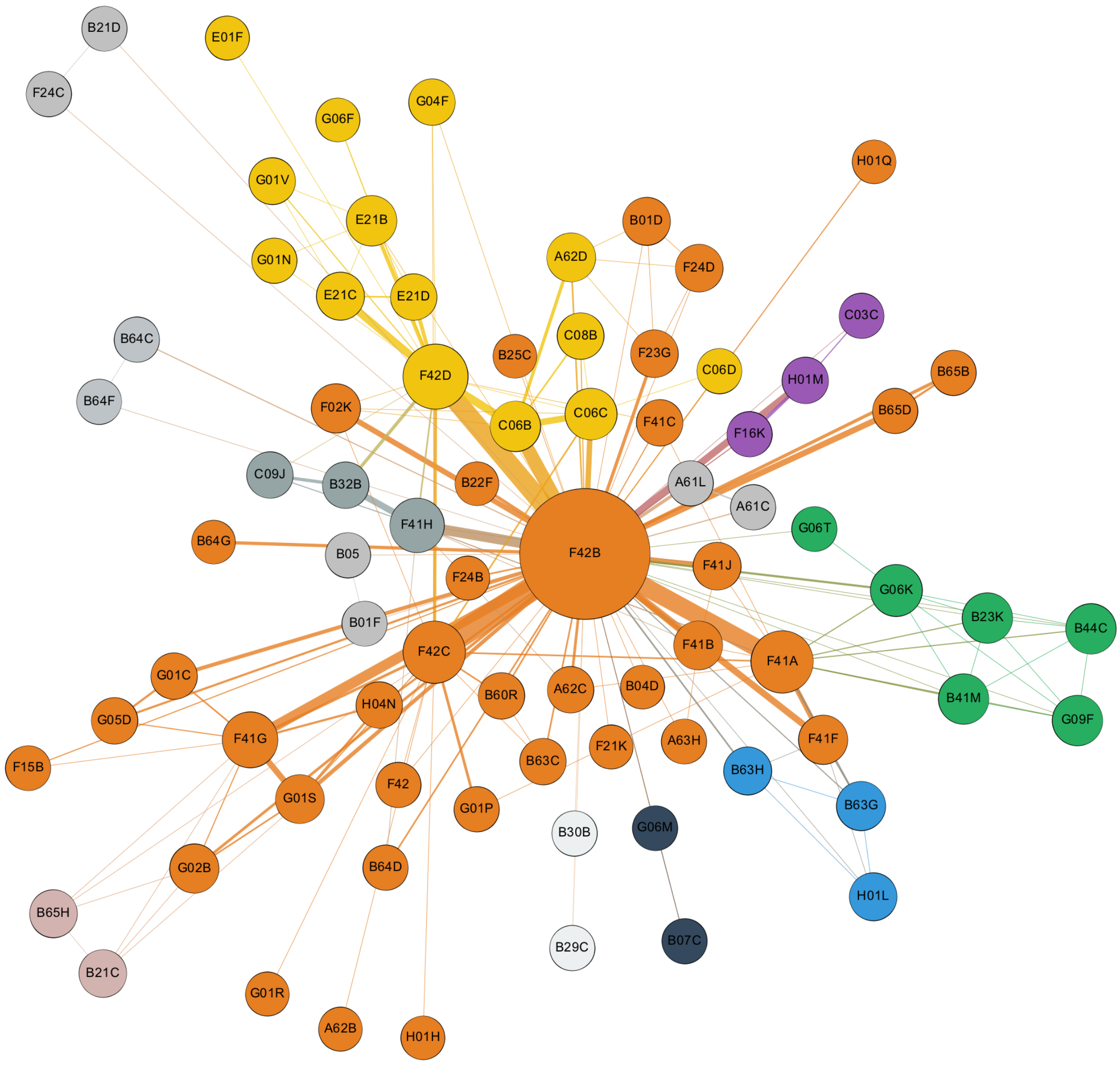

Figure 3.10 IPC Codes Graph-Network Fruchterman & Force Atlas Representation

Figure 3.10 is prepared with Fruchterman then Force Atlas algorithms and nodes are sized according to node's degree. It is seem that, nearly half of the nodes (orange), is gathered under the same modularity class. The key nodes are F41G and F41A. The following modularity class is yellow modluars. The significant node is F42D is here.

F41A and F42D nodes is discussed under "Centrality Measurements (Quantitative Analysis)" title. According to visualization above, F41A and F42D is a case for selecting which module.

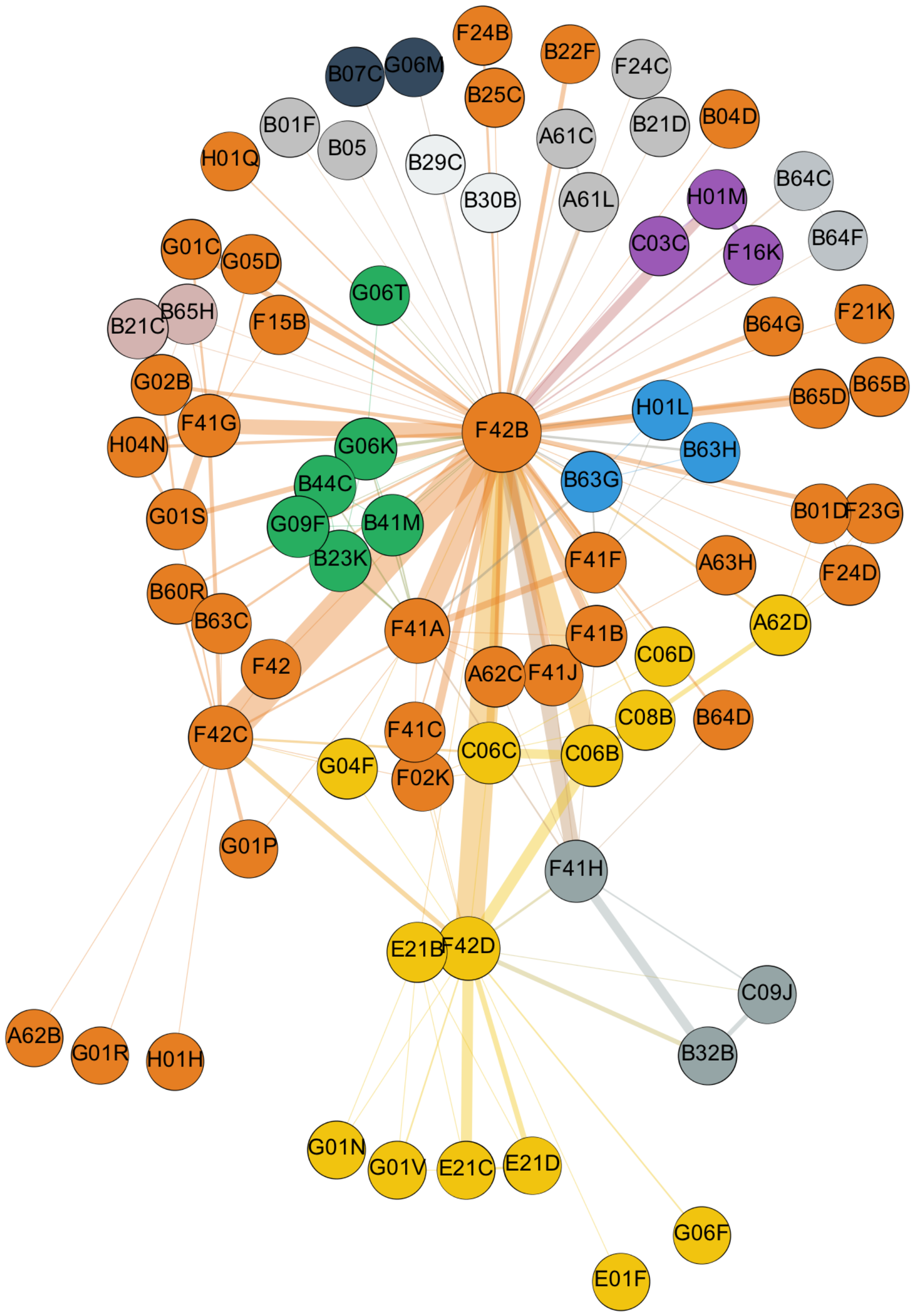

Figure 3.11 IPC Codes Graph-Network Yifan Hu (Proportional) Representation

Figure 3.11 is prepared with Yifan Hu algorithm and nodes sized same. Yifan Hu visualization has a grouping characteristic as position and distances of nodes imply for potential grouping. While F42 is a hub for yellow modularity Yifan Hu affirm that yellow modularity can be a single group, orange colored modularity group can be divided for different groups. Because orange nodes are distributed homogeneously from F42B center.

As a summary Figure 3.11 verify that, yellow modularity group with F42D hub, green modularity group with no hub, blue group with no hub, purple group with no hub should be taken as separate groups. For orange group, higher threshold application can compatible results between Yifan Hu and modularity.

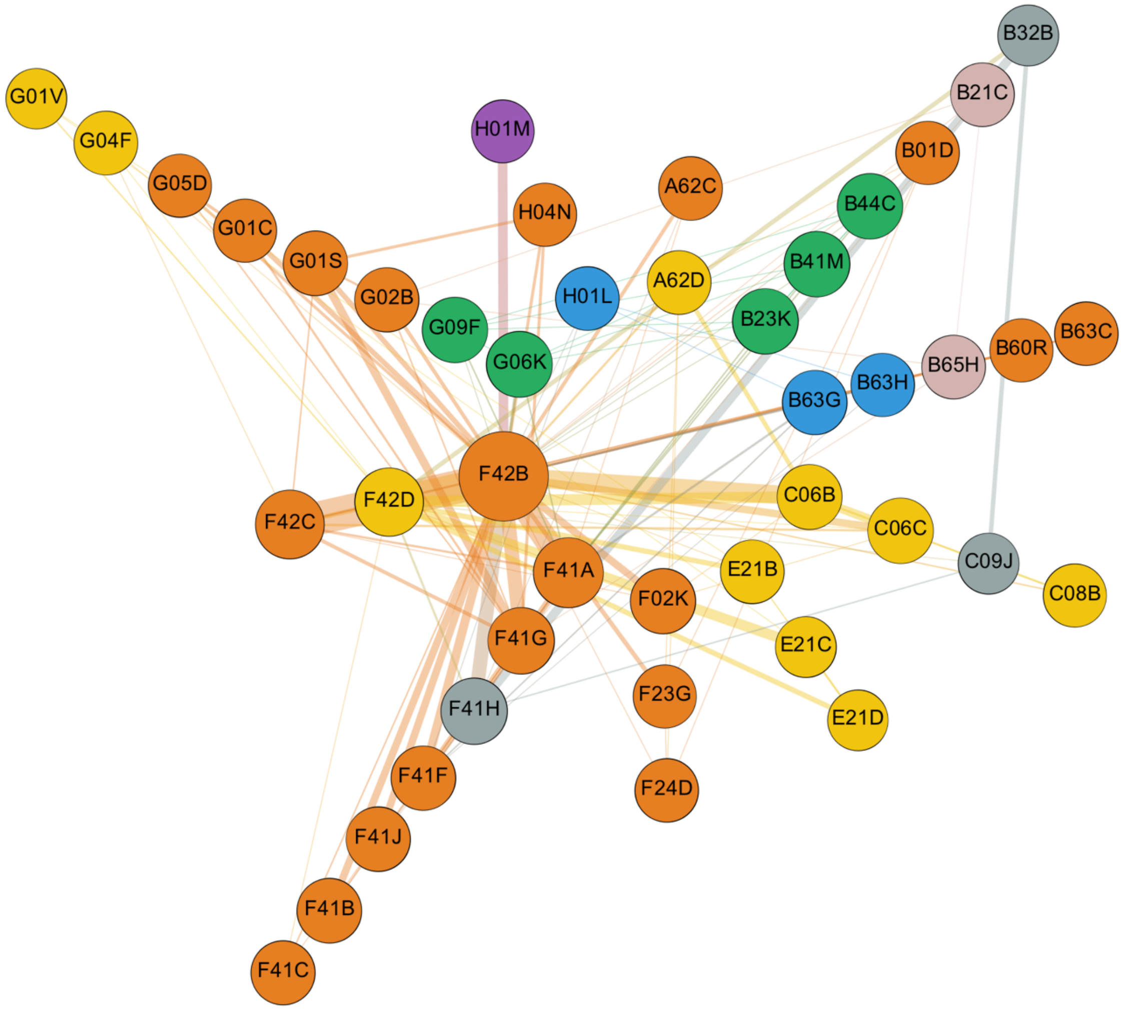

Figure 3.12 IPC Codes Graph-Network Yifan Hu (Proportional) Representation

Figure 3.12 shows relations between sections. Highest relations are seemed at F42-G[1] section relation then F42-C section relations. Minimum relation is seemed at F42-A[2] section relation.

---

[1] G section is the IPC symbol for Physics section which includes electricity, signaling, controlling and optics classes.

[2] A section is the IPC symbol for Human Necessities section which includes agricultural, foodstuff, personal and health classes.

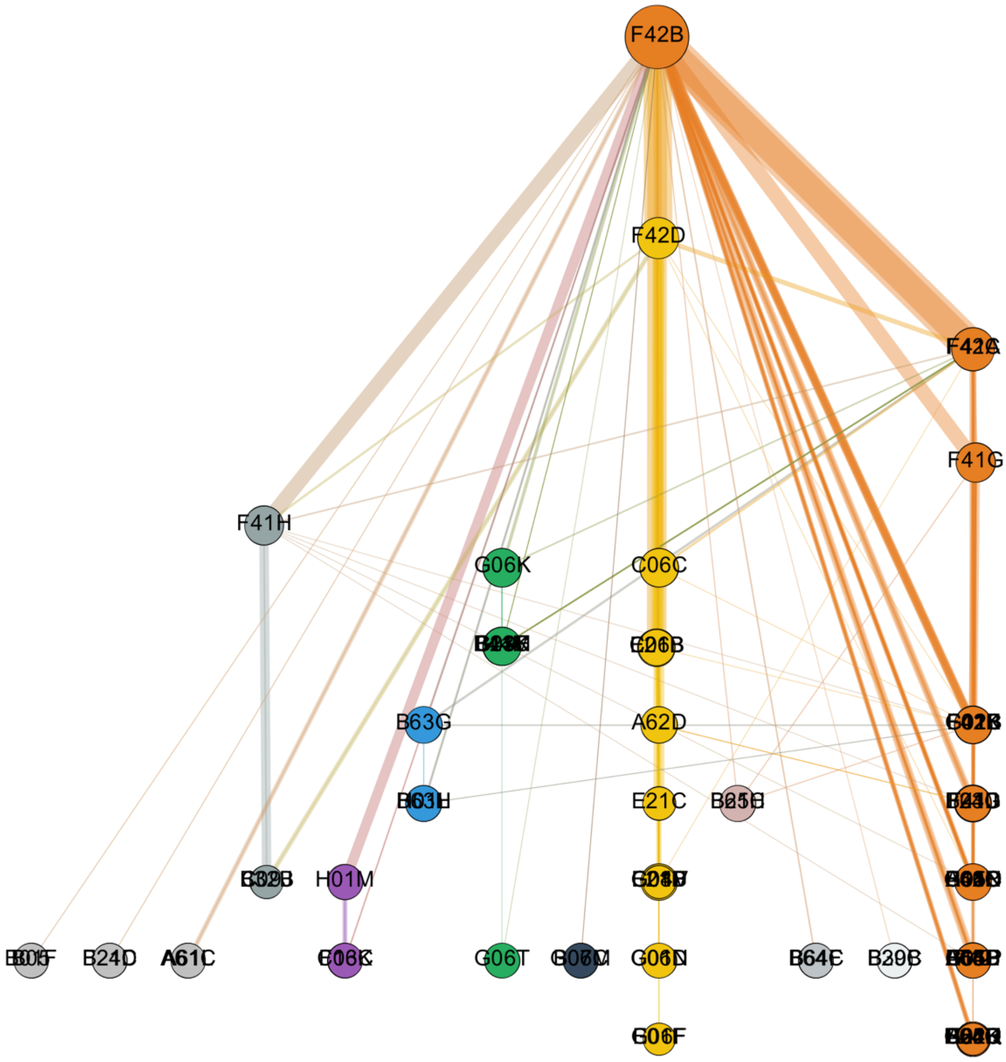

Figure 3.13 IPC Codes Graph-Network Representation as MDS Layout

A suitable technology group selection, by the terms of social network measures, should be in a modularity group and its degree average should be high. Figure 3.13 is prepared with MDS (Multidimensional Scaling) layout with horizontal axis as modularity classes and vertical axis as node's degree, up is for highest degree, bottom is for least.

According to this assumption first, lowest degrees should be eliminated like at Figure 3.14. First a level of degree should be filtered vertically (by degree) then low degree in-average modularity classes should be filtered. After filtering process, the remaining alternatives is stated at Table 3.7.

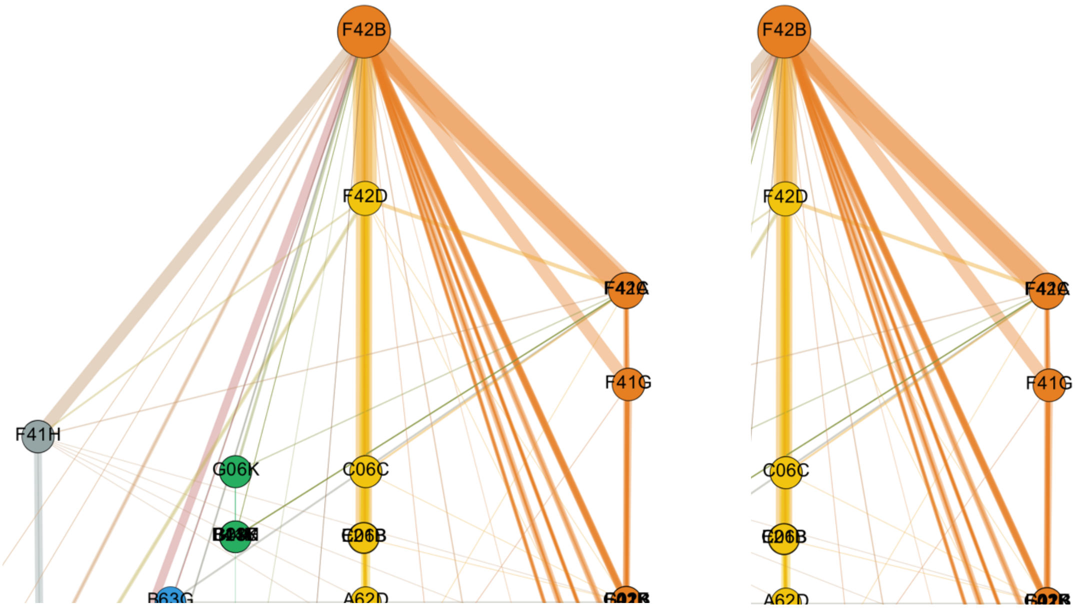

Figure 3.14 Lowest Degrees Eliminated from MDS Layout

Table 3.7 Technology Selection Alternatives

Option	**IPC Codes**	**Minimum Degree**	**Average Degree**
Orange	F42B, F41A, F42C, F41G	10	27
Yellow	F42B, F42D, C06C, C06B	6	24

Orange option, which categorized as ammunition, contains explosive charges (F42B), functional features to arms and ordnance (F41A), ammunition fuzes (F42C), weapon sights (F41G).

Yellow option, which categorized as blasting, contains explosive charges(F42B), blasting (F42D), detonating or priming devices (C06C), and explosive or thermic compositions (C06B).

CHAPTER 4

CONCLUSION

This work is subjected to analyze F42 patent category in Turkey to suggest alternatives for suitable technology selection. To select a suitable technology, the obvious way is to analyze relationship intellectual property registrations with the aspect of patent relationship.

Relationship analysis can be built via citation analyses which based on text mining techniques and machine learning or patent classification analyses like this example which can be analyzed by the power of little scripts. At this work, patent classification data is analyzed with Gephi application and complementary scripts.

Ammunition and Blasting (F42) technologies is diversified onto two categories. Like cited above at Table 3.7, for decision makers who prioritizes ammunition technology, orange alternative (F42B, F41A, F42C, F41G) is a suitable technology to track movements on this technology and to invest it. For decision makers who prioritizes blasting technology, yellow alternative (F42B, F42D, C06C, C06B) is a suitable. And for decision maker who has not a priority within ammunition and blasting, ammunition technologies[1] is more convenient to track and invest, by the conditions of registrations for Turkey by TPE.

[1] Details is handled at quantitative analysis section at page 27.

RESOURCES

[1] Kim Y. G., Suh J. H. and Park S. C., (2008) "Visualization of patent analysis for emerging technology", Expert Systems with Applications

[2] Yoon J., Park H.; Kim K., (2013) "Identifying technological competition trends for R&D planning using dynamic patent maps: SAO-based content analysis", Scientometrics

[3] Chan Y. S., Chen B. Y., (2011), "Utilizing patent analysis to explore the cooperative competition relationship of the two LED companies: Nichia and Osram", Technological Forecasting and Social Change

[4] World Intellectual Property Organization, "IPC Classification", http://www.wipo.int/classifications/ipc/en, 25 May 2018

[5] World Intellectual Property Organization, "Guide to the International Patent Classification", http://www.wipo.int/export/sites/www/classifications/ipc/en/guide/guide_ipc.pdf, 20 May 2018

[6] World Intellectual Property Organization, "Patent Map with Exercises by WIPS Co. Ltd", http://www.wipo.int/edocs/mdocs/sme/en/wipo_ip_bis_ge_03/wipo_ip_bis_ge_03_16-annex1.pdf, 15Apr 2018

[7] Sternitzke C., Bartkowski A., Schramm R., (2008) "Visualizing patent statistics by means of social network analysis tools", World Patent Information

[8] Franceschet M., “Degree Centrality”, https://www.sci.unich.it/~francesc/teaching/network/degree.html, Network Science, 09 Apr 2018

[9] Bisht J., “Closeness Centrality (Centrality Measure)”, https://www.geeksforgeeks.org/closeness-centrality-centrality-measure/, 12 Apr 2018

[10] Bisht J., “Betweenness Centrality (Centrality Measure)”, https://www.geeksforgeeks.org/betweenness-centrality-centrality-measure/, Jayant Bisht, 12 Apr 2018

[11] Bisht J., “Eigenvector Centrality (Centrality Measure)”, https://www.geeksforgeeks.org/eigenvector-centrality-centrality-measure/, 12 Apr 2018

[12] Du D., “Social Network Analysis: Centrality Measures”, http://www2.unb.ca/~ddu/6634/Lecture_notes/Lecture_4_centrality_measure.pdf, 13 Apr 2018

[13] WebHarvy, “What is Web Scraping?”, https://www.webharvy.com/articles/what-is-web-scraping.html, 01 May 2018

[14] Upwork, “Front-End Web Development: Client-Side Scripting”, https://www.upwork.com/hiring/development/how-scripting-languages-work/, 01 May 2018

[15] Lexisnexisip, “Importance of Name Normalization in Patent Research”, https://www.lexisnexisip.com/knowledge-center/importance-of-name-normalization-in-patent-research/, 04 May 2018

[16] Gephi, “About Gephi”, https://gephi.org/about/, 26 May 2018

[17] Ji X., Machiraju R., Ritter A., (2015), “Examining the Distribution, Modularity, and Community Structure in Article Networks for Systematic Reviews”, AMIA Annual Symposium Proceedings

[18] Li L., "Advanced Reading for the 2012 BlueTech Forum: Patent Landscape Analysis of M&A Activity in the Water Sector”, Cleantech Patent Edge, https://www.cleantechpatentedge.com/2012/05/analysis-of-ma-activity-in-the-water-sector/, 14 Apr 2018

[19] Trappey C. V., Wu H. Y., Taghaboni-Dutta F., Trappey A. J.C., (2009), “Using patent data for technology forecasting: China RFID patent analysis”, Advanced Engineering Informatics

[20] Jeong Y., Yoon B., (2014), “Development of patent roadmap based on technology roadmap by analyzing patterns of patent development”, Technovation

[21] Park S., Lee S. J. and Jun S. (2015), “A Network Analysis Model for Selecting Sustainable Technology”, Sustainability

[22] Disney A., “Social Network Analysis”, Cambridge Intelligence, https://cambridge-intelligence.com/keylines-faqs-social-network-analysis, 21 Apr 2018

APPENDIX-A

BETWEENNESS and EIGEN CENTALITY COMPARISON by KEYLINES, CAMBRIDGE INTELLIGENCE

Table 5.1 Comparison between Betweenness Centrality and Eigen Centrality [22]

	Betweenness Centrailty	**Eigen Centrality**
Definition	Betweenness centrality measures the number of times a node lies on the shortest path between other nodes.	Eigen centrality measures a node's influence based on the number of links it has to other nodes within the network, then goes a step further by also taking into account how well connected a node is, and how many links their connections have, and so on through the network.
What it tells us	This measure shows which nodes act as 'bridges' between nodes in a network. It does this by identifying all the shortest paths and then counting how many times each node falls on one.	By calculating the extended connections of a node, Eigen centrality can identify nodes with influence over the whole network, not just those directly connected to it.
When to use it	For finding the individuals who influence the flow around a system. Betweenness is useful for analyzing communication dynamics but should be used with care. A high betweenness count could indicate someone holds authority over, or controls collaboration between, disparate clusters in a network; or indicate they are on the periphery of both clusters.	Eigen centrality is a good 'all-round' SNA score, handy for understanding human social networks, but also for understanding networks like malware propagation.

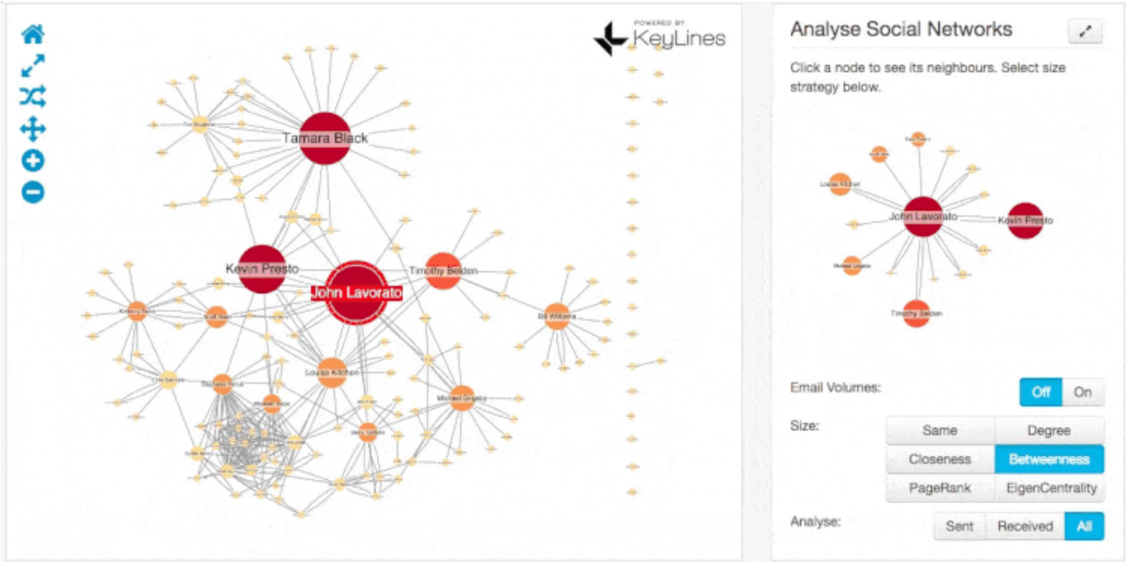

Figure 5.1 Nodes Sized by Betweenness Centrality [22]

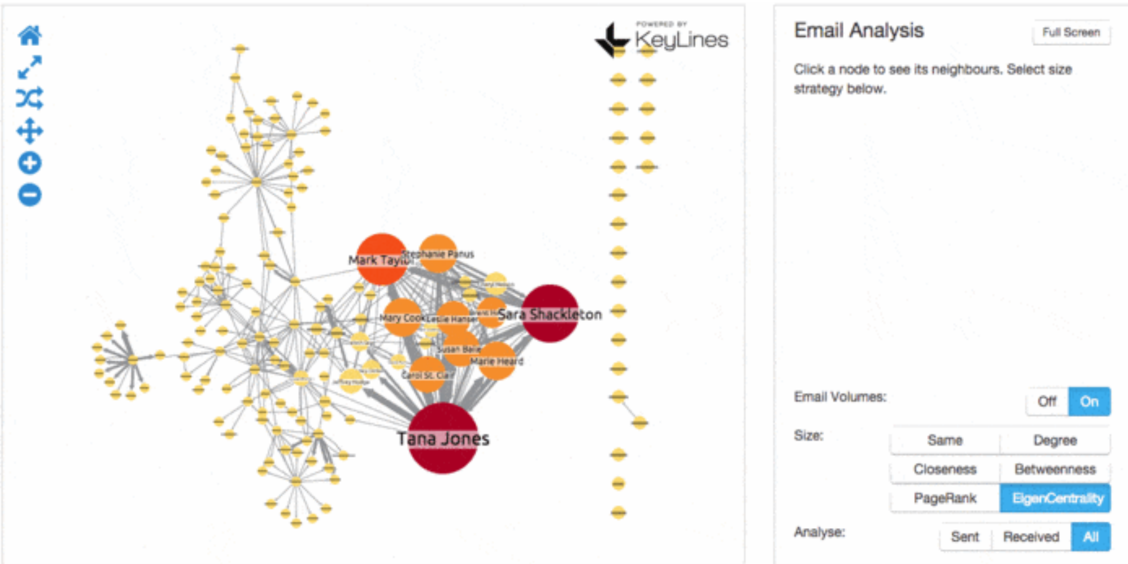

Figure 5.2 Nodes Sized by Eigen Centrality [22]