Building Textual Fuzzy Interpretive Structural Modeling to Analyze Factors of Student Mobility

Ronak Razavisousan (Corresponding author)
Information Systems Department
University of Maryland Baltimore County
Baltimore, USA
Ronak2@UMBC.edu

Karuna Pande Joshi
Information Systems Department
University of Maryland Baltimore County
Baltimore, USA
Karuna.Joshi@UMBC.edu

University of Maryland Baltimore County

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Abstract
There is a large amount of textual data present on web content that has the potential to answers many open questions in the field of humanities and human behavior. We have developed a novel methodology, called Textual Fuzzy Interpretive Structural Modeling (TFISM), that automatically analyses large textual datasets to identify the internal and external relationships between factors in student mobility. This methodology enhances approaches of Interpretive Structural Modeling (ISM) to allow the input type to be textual data. It is multi-disciplinary and integrates ISM with techniques from Artificial Intelligence, Text extraction, and information retrieval. We have validated this methodology on two different datasets from social media and academic articles. In this paper, we present the results of our study to identify the critical factors and most effective factors for global student mobility.

Keywords: Student Mobility, Fuzzy interpretive Structural modeling (FISM), Textual Fuzzy Interpretive Structural Modeling (TFISM)

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

1 Introduction

Student mobility is traditionally studied by economists and social scientists. Consequently, the motivation and pushing factors for studying this phenomenon are discussed from the economic or sociological point of view. Quantitative analysis and big data analysis can enable us to have a new look at this concept. In this research, we focused on the quantitative analysis of social media and examine how social media data can help the study of student mobility and how it can affect the process. We also tried to overcome the challenge of studying students' mobility from social media data.

Over half of the world population is currently online. A large volume of online data is generated by mobile users, social media applications, and sensors. Researchers increasingly recognize that these "digital traces" present an enormous opportunity to complement traditional sources of migration data and improve our knowledge of various migration aspects. Social media platforms capture the geo-location of their users and provide a useful dataset to study mobility patterns and immigration. Increased rates of human mobility all around the world [(United Nations, 2017), (Migration, 2019), (Castelli, 20 July 2018)] pushes the general population to recognize more geographical points than they have in the past. Traveling facilities have accelerated this phenomenon. For immigrants traveling to foreign locations, the internet provides data, knowledge, experiences, and concerns.

There are three roles involved with the concept of migration including “origin country”, “immigrant”, and “host country”. In the concept of student mobility, these three roles are functioning too. The origin refers to the location or country that migration starts. The host is the target country that a person arrives after leaving their home countries. An immigrant is a person who moves from the origin to the host. This research analyzed the barriers and motivations of student mobility/ migration, known in this field as pulling and pushing factors.

Researchers introduced market forces as one factor for the increasing rate of educated immigrants [(Gürüz, 2011)]. Some host countries such as Germany, Sweden, Switzerland, and France have plans to seek and attract more international students. These countries update the process of long-term residency for international students to attract them easily. Hence, highly educated immigrants will join the new population with fewer challenges and conflicts compared to people with low education. One of the examples of an enticing international student is equal tuition fees for them and open pathways to work opportunities [(Choudaha R. , 2017)]. European countries projected for the year 2025 to increase the hosted international students from 125 million to 263 million [(Gürüz, 2011)].

Using social media as a data source to study student mobility has challenges. Considering the growth of immigration and student mobility around the world pushes us to seek a sustainable source of data to study this subject. The very primary question to start the project is: Can the big data of social media serve as a basis for modeling human migration, specifically international student mobility? Another critical question is “How we can develop a model to understand the effectual factors of student mobility based on user-generated content?

Exposure to new cultures and providing sociological changes are considered positive effects of accepting immigrants. Therefore, understanding the flow of immigration and forecasting human behavior for mobility is valuable for societies, governments, and all participating parties [(Choudaha R. , 2017)].

Our proposed approach is inspired by one of the established decision-making and modeling method which is known as Interpretive Structural Modeling (ISM). We customized the ISM method to identify the critical factors for student mobility and extract a set of different related variables to structure into a comprehensive systematic model. It is invented by John Warfield in 1979 and it developed for a wide area of decision making and problem-solving [(Warfield J. , 1979)]. It helps to impose order and direction on the complexity of relationships among various elements. It is primarily intended as a group learning process, but individuals can also use it [(Attri, 2013)].

The uniqueness of the ISM is the creation of logical links between elements which helps to form a high-level visual map of the system or problem. The core competency of this method can be summarized into two futures first, simplicity in the sense of not requiring from the user i.e. viewpoint of advanced mathematical knowledge, and the second efficiency in terms of efficiency of computing time[(MSc., 2000)].

The input data for the ISM method is the knowledge of the domain experts. For identifying the interaction between two factors a person who has a deep understanding and expertise of domain values of the relationships. The main contribution of our work is considering the textual resources for each domain to find the relationships between factors. This novel integration enables us to easily enhance the decision-making process based on the noticeable amount of textual data. In this paper, we propose an approach to identify the factors and analyze the relationships and effectiveness of each factor.

2 Literature Review

Human migration is traditionally studied by economic researchers. There is a reach list of studies focused on human migration. Student mobility as a category of mobility and relocation can be studied under immigration researches. Student mobility overlaps with human migration and both these topics share factors. The definition of migration and legal definition of international students prohibits considering student mobility as a subsection of human migration.

2.1 Human migration and student mobility in economic and social sciences

This category contains all the reviewed student mobility studies from economic, sociological, psychological, cultural, demographical perspectives. Pulling and pushing factors are extracted from this set of articles. All these studies use a data source to identify the reason for student mobility and trends. None of these studies use social media data.

Verbik and Lasanowski (L Verbik, 2007) combined the official central government sources and extract the trend of the student movement. They categorized student mobility based on countries and economic status. In comparison to this article, we cannot categorize the countries based on the host or origin roles while we extract a list of highly effective countries. We also identify the power of impact and effectiveness for each country based on other factors. Interestingly the results are close to the former article. Another research project run by Rajika Bhandari (Bhandari, 2017) using project atlas as a data source. In this research, the author tried to identify the key factors of student mobility in higher education. The sample country was the United States and the results expanded to the world. In another interesting research, Choudaha and Kono (Choudaha R. a., 2012) identified the key factors of student mobility to guide U.S institutions for recruiting international students. The results from our research have a meaningful closeness to the result of this research. The same authors have other research to consider economic and political events to highlight the three waves of student mobility over time. Moreover, demographic changes cause student mobility patterns and influencing factors. These factors including economic, cultural, political, and sociological perspectives. The recent paper was published by Guruz (Gürüz, 2011).

The main difference between these researches and our project is the method that is engaged in the research. All the researches mentioned in this category used a Qualitative research method while in our research we applied a quantitative method for analyzing big data. We finally compare the extracted factors by the qualitative method with the factors extracted from quantitative methods.

2.2 Human Migration and Student Mobility in Social Media

The second group refers to the studies that used web data and social media to analyze the social factors of migration. Hughes et al. [(Hughes, 2016)] presented a wide range of data sources to study immigration in ways ranging from traditional methods to modern ones. They also provide a list of data sources and their features. We used the feature list of data sources to choose the most appropriate one. This research focuses on comparing the differences between traditional and modern methods. In our previous research [(R Razavisousan, 2019)], we use the results of the traditional method as the ground truth which is compared with the data from modern methods. McGregor et al. [(E. McGregor, 2013)] studied the feasibility of reviewing migration through social media. They studied the role of social media and its influences on the concept of human migration from four perspectives. Those four groups include 1) Influence of social media on migration, 2) migration integration and the role of social media in this process, 3) immigrant networks and the role of social media, 4) the role of social media in studying migration

[(E. McGregor, 2013)]. Our research is distinct from this work because we explicitly focus on student migration. Moreover, the core concept of their research highlights the role of social media in the field of migration. In contrast, we focus on the migration model and use social media as a data source to identify the power of each factor. While the results from our research confirm that social media and UGC can be a data source for studying human behavior and extract effective factors.

Researchers have also modeled human mobility. Hawelka et al. [(Bartosz Hawelka, 2014)] leveraged the geolocation of tweets to extract the mobility pattern. They also considered the seasonal timing and community network in this pattern. Another example is the study conducted by Pinto et al. [(A. L. Pinto, 2015)], which examined user behavioral activity and social media activity to justify migration between cities. They also consider the tweets' geolocations. Geo-located social media activity, such as posts on Twitter and LinkedIn, have been used to infer international migration flows [(Bartosz Hawelka, 2014)]. The main difference between these previous approaches and ours is in the type of data we analyzed. These studies have accessed the geolocation of users and identified the location changes in their approach, while we extract the pulling and pushing factors for movement based on the social media content shared by users.

Grover and Kar [(P. Grover, 2018)] extracted a practical solution for promotional marketing based on the social media activities of users. They built a Twitter engagement model for understanding user dynamics. Hong and Davison [(L. Hong, 2010)] studied topic modeling for microblogging datasets such as Twitter. McCallum et al. [(A. McCallum, 2016)] proposed a model to identify groups and topics among the text. Mikolov et al. [(T. Mikolov, 2013)] worked on the approaches to find a similarity between words, and they proposed a model to compute a vector for each word. Maas et al. [(A. L. Maas, 2011)] built a model that uses a mix of unsupervised and supervised techniques to learn word vectors that capture semantic terms. Our novel approach differs from these studies, as it combines topic modeling, clustering, and word embedding to determine the closeness of topics from different data sources. The main contribution of our work is to study student mobility based on analyzing UGC on Twitter and model the extracted factors by the invented method of textual fuzzy interpretive structural modeling (TFISM). We will explain the details of the proposed method TFISM and the results in the rest of the paper.

2.3 Methodology: Building a new model with Social Media data

Figure 1 presents the extracted critical factors for humans migrating based on economic and sociological researches. By looking deep at the motivation for student mobility we realized that international students have similar encouragements as people who migrate. Students consider the same factors regarding traveling overseas, but their priorities and the weight of factors vary.

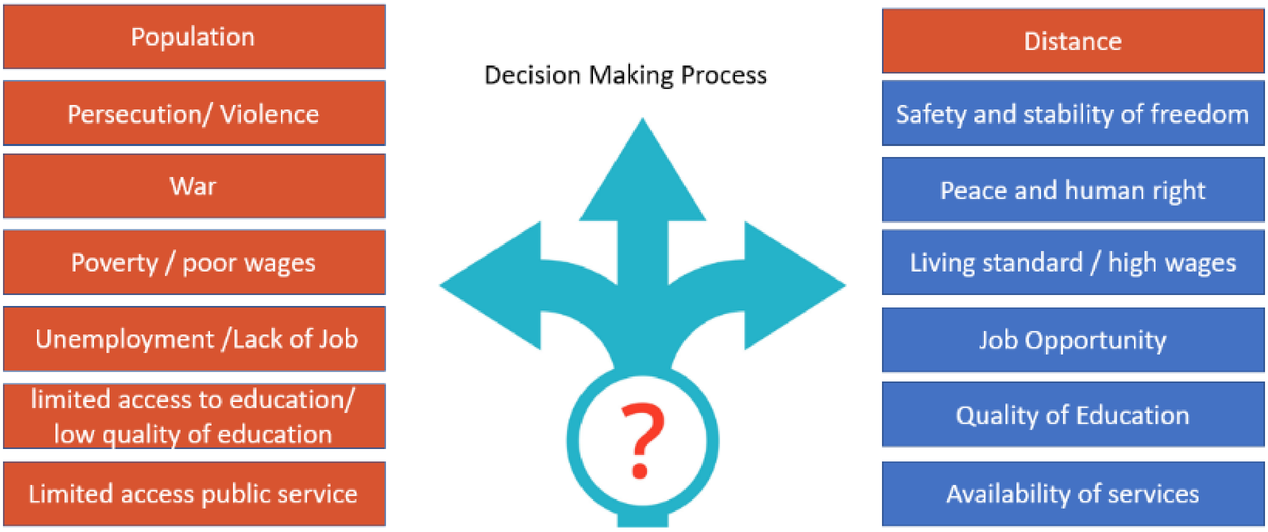

Figure 1: Pulling factor and pushing factors of student mobility. The red boxes show the pushing factors from origin countries and the blue boxes show the pulling factors from the host countries. The students decide about education and destination countries based on these factors

In other words, we can look at student mobility as a decision-making process and build a model for making this decision based on social media data. In this research, we try to identify the relationship between factors with the help of well-established modeling approaches. As is shown in Figure 1**Error! Reference source not found.**, several factors affect human migration. These factors are not very different for student mobility, while their weight for decision-making can be different. Interpretive Structural Modeling (ISM) is a methodology to identify the relationship between variables in the decision-making process or complex problem-solving. The designed method for process flow presents in Figure 2. We customized the methodology for modeling social media data. It is necessary to mention that ISM has an enhancement version which is called Fuzzy-ISM. In our proposed approach, we implement Fuzzy-ISM for using textual data from social media and automate this process.

ISM is an established and known methodology used in various fields. There are multiple extensions on the ISM method to adjust it for a wide range of domains. Fuzzy ISM is one of the extensions which we used to implement the ISM model based on textual data.

Fuzzy theory plays a significant role in dealing with vagueness and uncertainty in human language and thoughts in decision making. The assessment of decision-makers depends upon their past knowledge and experiences and often their estimations are articulated in equivocal linguistic terms. However, to integrate various opinions, experiences, ideas, and motivations of individual experts it becomes important to translate the linguistic judgments into fuzzy numbers.[(Zadeh, 1996)]

The integration of ISM with fuzzy sets provides flexibility to decision-makers to further understand the level of influences of one criterion over another, which was earlier present only in the form of binary (0,1) numbers[(Khatwani, 2015)]. 0 represents no influence and 1 represents influence. Due to this, the decision-maker is left with only the option of saying 0 or 1 irrespective of the level of influence whether it is low, high, or very high. The proposed Fuzzy-ISM approach takes care of this issue and gives wider flexibility to express the level of influence using fuzzy numbers. Fuzzy Interpretive Structural Modeling (FISM) is a better match for our purpose. In this approach instead of a binary point of view to identify the relationship between variables, the relationships are ranked based on a range of numeric values between 0 and 1 such as 0, 0.1, 0.3, 0.5, 0.7, 0.9, and 1.

ISM and its extensions applied to the various domain that we look at some of them. For instance, Shrimali [(Shrimali, 2019)] applied the ISM model to the manufacturing domain to extract factors that affect the successful implementation of lean manufacturing; also identified the interaction between factors. For the construction domain, Sandbhor and Botre [(Sayali Sandbhor, 2014)] used TISM to study affecting factors on labor productivity to enhance the productivity of the construction projects and improve associating factors. Another domain addressed by the ISM method is seismology. Ahmad, Mahmood et.al [(Ahmad, 2019)]used the ISM method to highlight the soil factors and the relationships with earthquakes. Besides, e-commerce took advantage of Fuzzy ISM. Valmohammadi and Dashti [(Valmohammadi, 2016)] applied FISM to identified and highlight inherent interactions among these barriers of e-commerce. For the technology domain, the ISM integrate by NPD New Product Development (NPD) to evaluate various technologies for the new product [(Lee, 2011)].

The major adaptation in the original ISM version that helps us to solve the problem, happens in two first steps. Before introducing the next step, these two steps with new updates have been explained. Having access to the knowledge of experts for a specific domain is the key principle of ISM methodology. Without the knowledge of the specialist, the first step of ISM methodology will not develop. Moreover, finding the expert of the domain who is familiar with the concept of ISM will make the process harder.

The main contribution of our research is customizing the process in a way that helps to build the SSIM automatically and without the knowledge of experts. We consider two different data sources including social media data and academic articles.

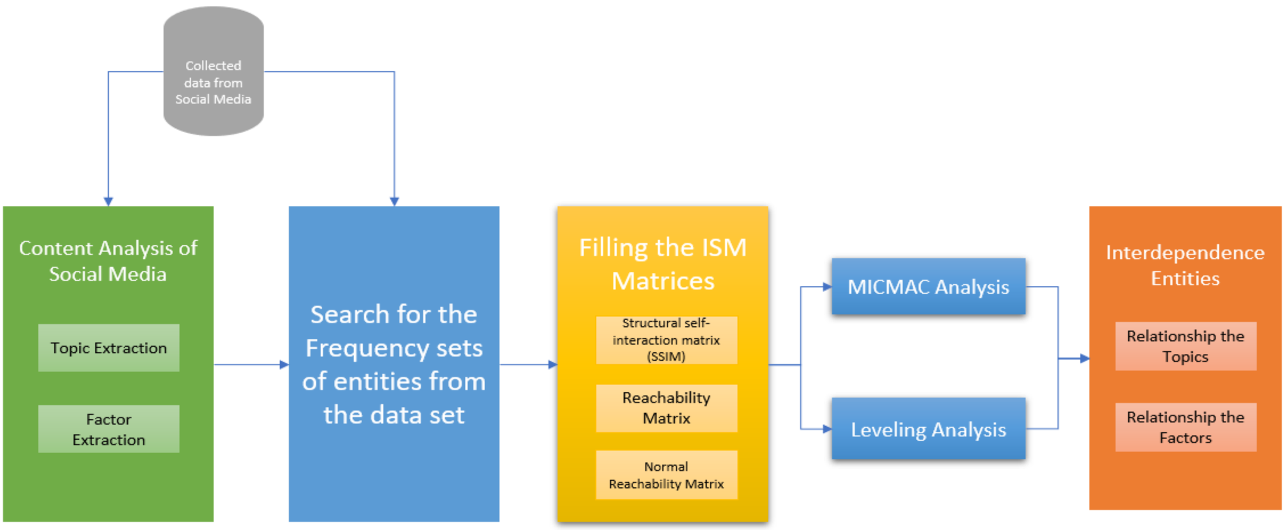

Figure 2: Methodology of implementing ISM and Process Flow

Our goal is to implement a process flow that enables us to build SSIM automatically instead of expert knowledge. In other words, the SSIM in the new approach is evolved based on the content of social media or academic articles (extracted from big data). The process flow of creating SSIM based on the type of data has some changes. The process flow of implementing Fuzzy-ISM for two data types explain separately and the quality of results and the method's validity will discuss in the next part.

3 Experiment

3.1 Data Collection

We used two sets of datasets in this research. Social media data is a collection of tweets related to student mobility that contains more than 7 million tweets and a list of academic articles focused on topics of student mobility. This list carries sixteen articles. The process of building these datasets refers to another research run by the same authors [(R Razavisousan, 2019)].

3.2 Process flow of using Fuzzy- ISM in data modeling of social media

We have a dataset of Twitter posts that is collected based on the student mobility keyword list. Dataset has more than seven million records. In this chapter, we try to run the fuzzy-ISM for the student mobility data. The flow diagram of Fuzzy ISM for social media data presented in Figure 3, and for simplifying the process explaining the steps are designed like the Warfield-ISM version. Gray-colored items in figure 3 are the parts designed for TFISM implementation. These parts are the main contribution of the research. The blue items in figure 3 represent the original steps of the ISM method which are customized to apply to the new approach.

3.2.1 Step 1: Create Structural Self-Interaction Matrix (SSIM) based on the extracted topic from social media

Extracting the variable automatically is one of the ultimate goals in this approach, for this purpose we decide not to apply the opinion of experts in the field of student mobility and enhance the method by extracting the key factors of content from social media. Using the extracted topic from topic modeling techniques is part of our solutions. Consequently, we will reach a list of topics and use them as variables to build the SSIM. In this process, we take advantage of the Latinate Dirichlet Allocation (LDA) technique for extracting a list of topics.

For running LDA, two different settings were tried. First, having a high number of clusters (for example 8) and a smaller number of topics in each cluster (for example 5); for the second set, having a low number cluster (for example 2) and a high number of topics in each cluster (for example 20). Prototyping the methods with sample data shows that the second set provides better results.

3.2.2 Step 2: Filling the SSIM automatically based on social media data

An important enhancement in the implementation of the Fuzzy ISM methodology for social media data is filling the SSIM by an automated process. In other words, there is less employment of knowledge of experts, while the benefit of general knowledge or large population in social media is adapted. The frequency of appearance-specific topics considered for identifying a relationship between two topics of i and j. The number of times that topic i and j come in each reference document will put in the cross-section cell of i and j in SSIM., in Figure 4, the highlighted cell which has the value of 2 shows the number of times "policy" appears when "issue" is already showing up. For filling the ISSM, the topic in each row consider as the main topic, then the appearance of the topic in columns will enumerate. Besides, the values in gray cells show the total number that the rows' topic occurred. For explanation, the topic "issue" was repeated 6 times in all sample tweets, and it comes with a policy two times. It is not hard to find out that the SSIM is a diagonal matrix because display two words in the dataset are not associated with the column and row; the appearance of "issue" with "policy" similar to the appearance of "policy" with "issue". It is good to know that the SSIM in an original version also has the same characteristics.

For having a full-filled SSIM the value of the filled cells copied symmetrically based on the diameter. Figure 4 shows the final SSIM matrix filled with sample Twitter data. The major difference between the original version of fuzzy-ISM and adjusted semi-automated versions is in the type of cells' values. Formerly, the SSIM contains the value of (V, A, X, and O) which are converted to 0 and 1 (0,1) and for fuzzy extension converted to the rational number between 0 and 1 [0,1], while in this approach SSIM can be filed with any whole numbers as presented in Figure 4.

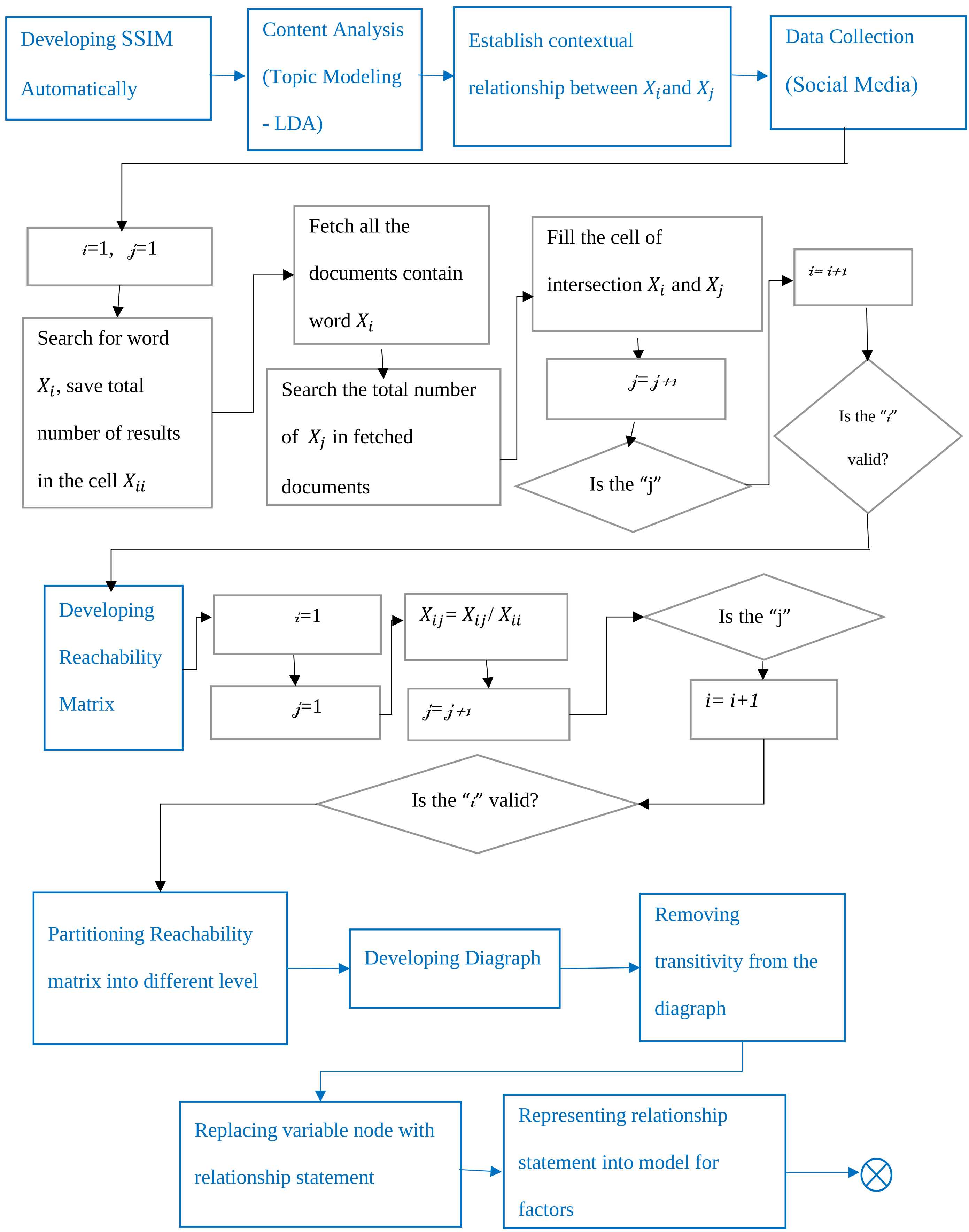

Figure 3: Process flow of Fuzzy-ISM for textual data

		schools	family	Internationalstudent	History	education	University	work	Canada	International	policy	Gvernment	created	week	around	issues	House	Speaker	UK	Laredo	border	Nancy Pelosi	security	discusses	TAMIU	US
		25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
US	1	0	0	0	0	2	4	2	0	0	0	4	4	0	2	1	0	0	2	0	0	0	0	0	0	15
TAMIU	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	8	0	8	8	8	8	8	8	0
discusses	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	8	0	8	8	8	8	10	8	0
security	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	8	0	8	8	8	8	8	8	0
Nancy Pelosi	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	8	0	8	8	8	8	8	8	0
border	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	8	0	8	8	8	8	8	8	0
Laredo	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	8	0	8	8	8	8	8	8	0
UK	8	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	2
Speaker	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	10	0	8	8	8	8	8	8	0
House	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	6	0	6	6	6	6	6	6	0
issues	11	3	0	0	0	0	0	0	0	0	2	0	0	0	2	6	0	0	0	0	0	0	0	0	0	1
around	12	2	0	0	0	0	0	2	0	0	2	0	0	0	8	2	0	0	0	0	0	0	0	0	0	2
week	13	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0
created	14	0	0	0	1	1	4	1	0	0	0	4	6	0	0	0	0	0	0	0	0	0	0	0	0	4
government	15	0	0	0	0	1	4	0	0	0	0	5	4	0	0	0	0	0	0	0	0	0	0	0	0	4
policy	16	3	0	0	0	0	0	1	0	0	5	0	0	0	2	2	0	0	0	0	0	0	0	0	0	0
International	17	1	0	3	0	1	1	0	3	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Canada	18	0	0	2	0	2	0	0	6	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
work	19	1	2	0	0	0	0	12	0	0	1	0	1	0	2	0	0	0	0	0	0	0	0	0	0	2
University	20	0	0	0	0	1	10	0	0	1	0	4	4	0	0	0	0	0	3	0	0	0	0	0	0	4
education	21	0	0	1	1	13	1	0	2	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	2
History	22	0	0	0	4	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
internationalstudent	23	0	0	3	0	1	0	0	2	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
family	24	0	3	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
schools	25	11	0	0	0	0	0	1	0	1	3	0	0	0	2	3	0	0	0	0	0	0	0	0	0	0

Figure 4: sample of full filled SSIM by Fuzzy ISM for social media

3.2.3 Step 3: Building the reachability matrix and normalizing the reachability matrix

In the primary version of ISM, the reachability matrix is a numeric matrix that shows the dependency and driving power between i and j variables. Each value shows how i depend on j and how i can drive j. For extracting this concept from the numeric data, the values in each row are divided by the total number of topic appearances which is shown in dark-gray cells in Figure 4.

3.2.3.1 How to extract driving power and dependency?

At this stage, the main burden is interpreting the driving power and dependency from the collected data. The simple concept that can be extracted from prior versions of ISM is the definition and relationship between driving power and dependency. In other words, the relationship of variable i with other variables identifies driving power and dependency. Figure 5**Error! Reference source not found.** shows the association between these two concepts.

As demonstrated in Figure 5**Error! Reference source not found.** dependency of i to j defines as a dependency, and dependency j to i, illustrate the driving power of i and j. Therefore, we can define dependency and driving power for a single variable based on the dependency of the index variable to others and the dependency of other variables to the index variable. With the help of this explanation, we need to define the two-way dependency of the variable. Two-way dependency means how i depend on j and how j depends on i. In the presented case, finding the dependency of two variables from the appearances in the text is the goal. The proposed solution is based on the simultaneous appearance of two variables in the collected data. We know that the number of times i appear with j is equal to the number of j appears with i, this number calls by X. The total number that i is showed up named I; and the total number that j is coming up by J.

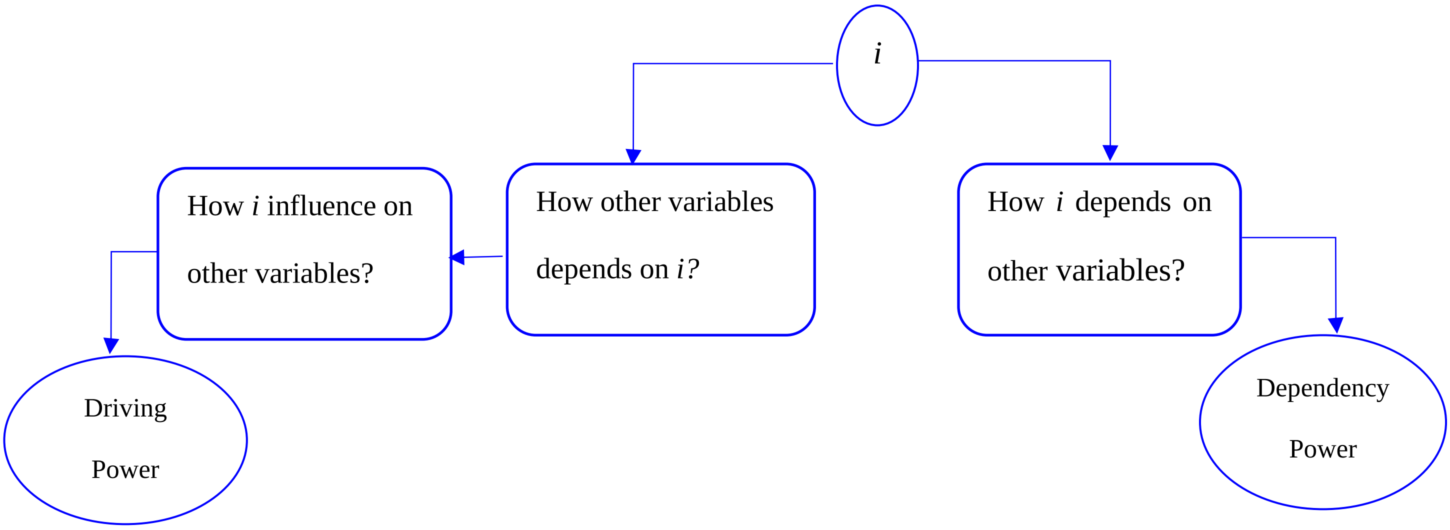

Figure 5: Variables associations - Driving Power and Dependency

Based on our definition, X dividing by I; (X/I) presents dependency j to i, which can be conveyed as the driving power of i. Moreover, the result of X dividing by J; (X/J) shows the dependency i to j. which means the dependency of i. Table 1 presents a hypothetical example with only two variables, i and j. The value for I is 10 and J is 50. The number of appearances of i and j together is 10. It means from 10 times i show up always is followed by j. It is supporting the driving power of i. In this case driving power for i calculated 1, which is the maximum rate for driving power. Since the 50 times j appears only 10 times come with i that means dependency of i is 0.2. Consequently, the value in rows shows the driving power, and the values in columns interpret as a dependency power.

Table 1: Sample for explaining the concept of Dependency and Drovong Power

A

	j	**i**	
i	10	10	1
j	50	10	0.2

Divide by blue cell in row

B

	j	**i**	
i	1	1	1
j	1	0.2	0.2

4 Result and Validations

We implemented a series of updates and changes in the original version of ISM to customize the process for social media data. Based on these adjustments, we must update the validation methodology and use other validation techniques to evaluate the reliability of our proposed solution. The following three types of validation are considered.

1. Association rule mining
2. Basic structural models are validated using standard logic and mathematics.
3. Interpretive structural models are validated using expert domain knowledge.

each method will be discussed in detail the high-level view of the proposed method present in Figure 6.

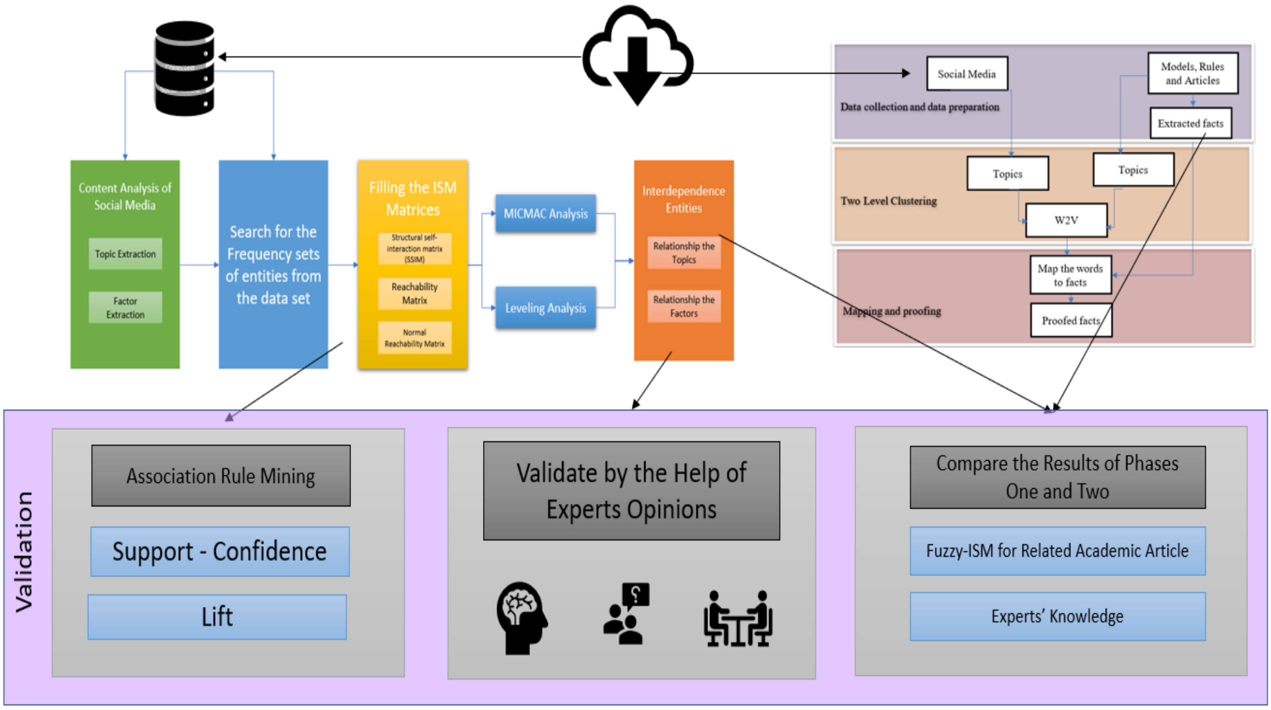

Figure 6:Validation Methodology

4.1 Association Rule mining:

In this case, the possible relationship between sets of two topics is validated. The frequency of two words is searched in the whole Twitter data sets with 7 million tweets. With the help of the frequency of the topics, Figure 7 can present the support values for each set of words or for every single topic in diameter cells which are highlighted by green color. The value for lift is calculated for topics and presents in Figure 8Figure 8 and all the values greater than 1 shows by red color.

For calculating the support of each factor, first, we count the number of the appearance of each set of two topics in the dataset. Each row contains the repetition of the topic in the row with each topic in the column. The value in the diameter cells demonstrates the summation of all the values in the rows. Each value in the diameter cells is divided by the total number of topic sets and gave us the support values (total number of topic sets reached by the summations of all the diameters values). Topics with the highest support are the "border", "illegal", "law", "OPT", and "Visa".

Suport	T=206	visa	cpt	opt	school	Pathway	family	student	History	education	University	work	World	Canada	International	policy	government	law	week	sexual	cost	security	House Speaker	UK	illegal	border	financial	global	Trump	US
		29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
US	1	0.003	0.000	0.001	0.000	0.000	0.000	0.003	0.000	0.001	0.006	0.005	0.001	0.001	0.001	0.002	0.002	0.002	0.000	0.000	0.002	0.000	0.000	0.001	0.005	0.005	0.000	0.002	0.013	0.102
Trump	2	0.003	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.012	0.001	0.000	0.000	0.004	0.001	0.002	0.001	0.000	0.009	0.001	0.001	0.000	0.016	0.013	0.000	0.002	0.153	
global	3	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.012		
financial	4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.005			
border	5	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.012	0.002	0.000	0.000	0.009	0.002	0.000	0.007	0.077				
illegal	6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.000	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.008	0.001	0.002	0.000	0.073					
UK	7	0.001	0.000	0.001	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.003	0.001	0.001	0.000	0.000	0.001	0.000	0.000	0.033						
House Spe	8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.011							
security	9	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.013								
cost	10	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.025									
sexual	11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.007										
week	12	0.003	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.017											
law	13	0.001	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.000	0.000	0.001	0.068												
governme	14	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.002	0.001	0.000	0.000	0.000	0.000	0.015													
policy	15	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.018														
Internatio	16	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.000	0.001	0.000	0.000	0.000	0.000	0.013															
Canada	17	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.023																
World	18	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.006																	
work	19	0.003	0.000	0.002	0.001	0.000	0.001	0.001	0.000	0.000	0.001	0.071																		
University	20	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.040																			
education	21	0.000	0.000	0.001	0.001	0.000	0.000	0.001	0.000	0.008																				
History	22	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001																					
student	23	0.001	0.000	0.000	0.001	0.000	0.000	0.030																						
family	24	0.000	0.000	0.000	0.000	0.000	0.013																							
Pathway	25	0.000	0.000	0.000	0.000	0.000																								
schools	26	0.000	0.000	0.001	0.012																									
opt	27	0.000	0.000	0.076																										
cpt	28	0.000	0.024																											
visa	29	0.053																												

Figure 7: Support for sets of topics

With the same logic, the confidence and lift were calculated for the topics. The red cells in Figure 8 are the topics with the lift greater than 1. The topics which have the highest lift value in the row are “The US”, “Trump”, “border”, “illegal”, “international student”, and “University”.

The results of association rule mining show the expected values and the relations between factors are very close to the previous results that we expected. In the other part of the validation phase, we can refer to these results and compare them with other methods’ outcomes.

lift	T=206	visa	cpt	opt	school	Pathway	family	student	History	education	University	work	World	Canada	International	policy	government	law	week	sexual	cost	security	House Speaker	UK	illegal	border	financial	global	Trump	US
		29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
US	1	0.644	0.047	0.088	0.412	1.824	0.293	0.988	3.157	1.129	1.422	0.681	0.805	0.380	0.634	1.050	1.454	0.354	0.233	0.000	0.942	0.117	0.388	0.406	0.738	0.610	0.140	1.371	0.808	
Trump	2	0.469	0.003	0.086	0.032	0.000	0.850	0.091	1.590	0.138	0.106	1.622	1.392	0.102	0.016	2.410	0.674	0.343	0.609	0.227	3.547	1.038	1.232	0.025	2.187	1.717	0.080	1.827		
global	3	0.018	0.001	0.097	0.040	0.000	0.000	0.036	0.000	0.254	0.008	0.011	1.143	0.012	0.099	0.012	0.000	0.006	0.009	0.009	0.004	0.602	0.000	0.022	0.006	0.100	0.013			
financial	4	0.001	0.000	0.002	0.005	0.000	0.005	0.014	0.000	0.007	0.006	0.002	0.005	0.011	0.007	0.002	0.004	0.000	0.009	0.005	0.010	0.000	0.003	0.000	0.008	0.000				
border	5	0.022	0.000	0.128	0.024	0.000	0.099	0.061	0.047	0.004	0.004	0.063	0.040	0.072	0.016	0.385	0.050	1.766	1.022	0.000	0.048	6.765	1.797	0.013	0.929					
illegal	6	0.072	0.000	0.035	0.103	0.000	0.131	0.008	0.071	0.045	0.002	0.866	0.060	0.064	0.007	0.111	0.307	0.674	0.020	0.019	3.085	0.863	1.932	0.057						
UK	7	0.095	0.020	0.106	0.055	0.365	0.503	0.162	0.142	0.295	0.110	0.145	0.075	0.076	0.233	1.415	0.473	0.114	0.098	0.090	0.302	0.022	0.011							
House Spe	8	0.002	0.000	0.004	0.003	0.000	0.050	0.029	0.047	0.000	0.010	0.019	0.005	0.001	0.012	0.009	0.008	0.024	0.007	0.000	0.062	0.078								
security	9	0.008	0.000	0.110	0.003	0.000	0.005	0.019	0.000	0.019	0.002	0.010	0.482	0.000	0.037	0.040	0.017	0.013	0.870	0.000	0.003									
cost	10	0.286	0.010	0.032	0.053	0.365	0.084	0.026	0.000	0.101	0.003	0.431	0.005	0.032	0.051	0.067	0.544	0.042	0.007	0.000										
sexual	11	0.005	0.000	0.060	0.098	0.000	0.002	0.006	0.000	0.127	0.008	0.007	0.000	0.003	0.005	0.063	0.529	0.002	0.063											
week	12	0.475	0.035	0.028	0.166	0.000	0.042	0.188	0.190	0.093	0.176	0.080	0.020	0.023	0.115	0.044	0.035	0.056												
law	13	0.120	0.019	0.100	0.544	0.000	0.233	0.110	0.142	0.064	0.042	0.647	0.477	0.052	0.032	0.257	0.405													
governme	14	0.022	0.000	0.024	0.016	0.000	0.144	0.612	0.071	0.105	0.379	0.131	0.005	0.124	0.053	0.190														
policy	15	0.012	0.001	0.096	0.000	0.000	0.074	0.036	0.024	0.131	0.000	0.136	0.030	0.068	0.062															
Internatio	16	0.019	0.001	0.006	0.354	0.000	0.156	2.096	0.095	1.215	0.029	0.047	0.413	0.077																
Canada	17	0.109	0.000	0.015	0.016	0.365	0.183	0.161	0.000	0.045	0.039	0.279	0.060																	
World	18	0.058	0.008	0.015	0.000	0.000	0.002	0.012	0.119	0.019	0.048	0.019																		
work	19	0.545	0.026	0.276	0.541	0.000	0.731	0.381	0.926	0.258	0.192																			
University	20	0.001	0.000	0.009	0.399	0.365	0.050	0.474	1.044	0.172																				
education	21	0.006	0.026	0.086	0.533	0.365	0.020	0.198	0.000																					
History	22	0.001	0.000	0.004	0.005	0.000	0.002	0.008																						
student	23	0.099	0.004	0.035	0.892	0.000	0.159																							
family	24	0.091	0.010	0.011	0.016	0.000																								
Pathway	25	0.000	0.000	0.000	0.000																									
schools	26	0.012	0.001	0.074																										
opt	27	0.024	0.006																											
cpt	28	0.001																												
visa	29																													

Figure 8: Lift for sets of topics

4.2 Validating phase by comparing social media and academic articles

For this comparison, in addition to our collected dataset from Twitter, we use collections of sixteen academic articles that mainly focus on student mobility over the world. The list of these articles comes from another research.[(R Razavisousan, 2019)]

4.2.1 Fuzzy- ISM for the related academic articles

For comparing two sets of results, we must use the same variables and indicators. In previous research, we use the topics from Twitter and articles, then find the similarities between these two sets of topics. At this point, we use a unique set of topics and run Fuzzy-ISM based on the frequency of these topics in two datasets. The topics are the most frequent topics from Twitter and articles. The topic list contains 76 topics that are presented in Table 2.

Table 2: Frequent topics in Tweeter and article datasets

"abroad"	"class"	"family"	"home"	"law"	"performance"	"security"	"urban"
"academic"	"cost"	"fee"	"host"	"level"	"policy"	"service"	"US"
"achievement"	"culture"	"foreigner"	"Income"	"link"	"population"	"social"	"utilization"
"Australia"	"degree"	"France"	"India"	"living cost"	"poverty"	"student"	"Vietnam"
"Family background"	"destination"	"GDP"	"International"	"mobile"	"quality"	"tuition"	"work"
"Bologna"	"diploma"	"gender"	"Italy"	"media"	"relatives"	"Turkey"	"world"
"border"	"economic"	"Germany"	"Job"	"network'	"resource"	"UK"	
"Brazil"	"education"	"global"	"Kazakhstan"	"origin"	"rural"	"undergraduate programs"	
"Canada"	"enrollment"	"government"	"knowledge"	"overseas"	"Saudi "	"unemployment"	

"China"	"Europe"	"growth"	"language"	"parents"	"school"	"university"	

Figure 9 presents the topic distributions for the articles and Figure 10 presents the topic distribution for the tweets. These plots are built based on MICMAC analysis. As shown in both Figure 9 and 10 considerable number of topics has low dependency power and driving power, so they are located at the left-bottom of the plots. In Figure 9, “foreign”, “student”, “US” and “work” are the outliers and in Figure 10, “foreign”, “Family”, and “student” are the outliers. MICMAC analysis outliers indicated the variables with high dependency power or driving power or both. It is not surprising that the words, “foreign”, “student” are in the outliers’ range in both articles and tweets dataset.

For having a clear view of the topics’ dependency and driving power we remove the outliers from the plot and the rest of the topics are shown in wider space. Figure 11 and Figure 12 present topic distributions without the outliers. In Figure 13 both data are presented in one plot that gives us a view of distributions of topics from two datasets. The blue dots are presenting Twitter’s topic and the orange dots are presented articles.

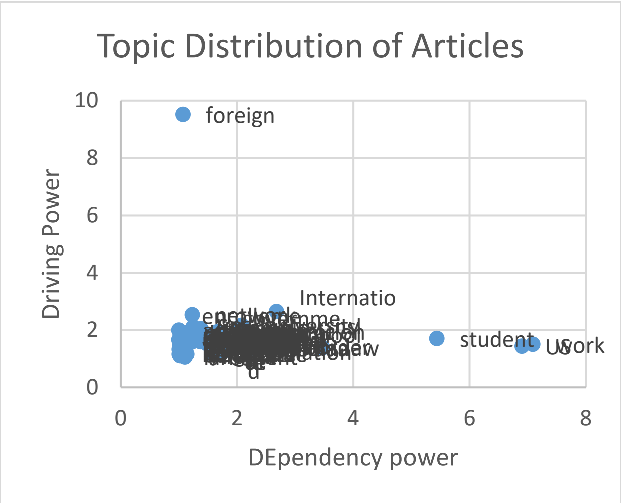

Figure 9: Topic distribution of articles with outliers

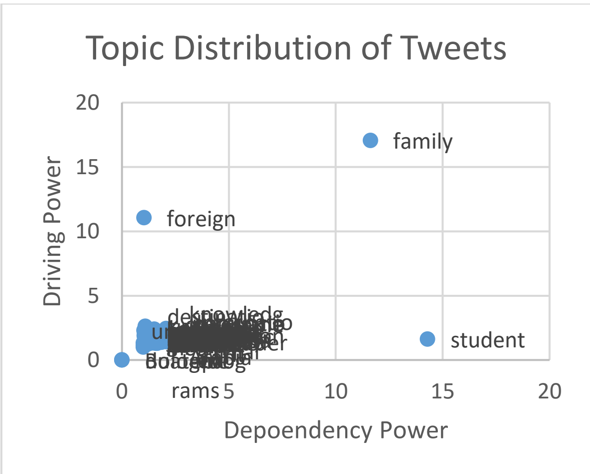

Figure 10: Topic distribution of Tweets with outliers

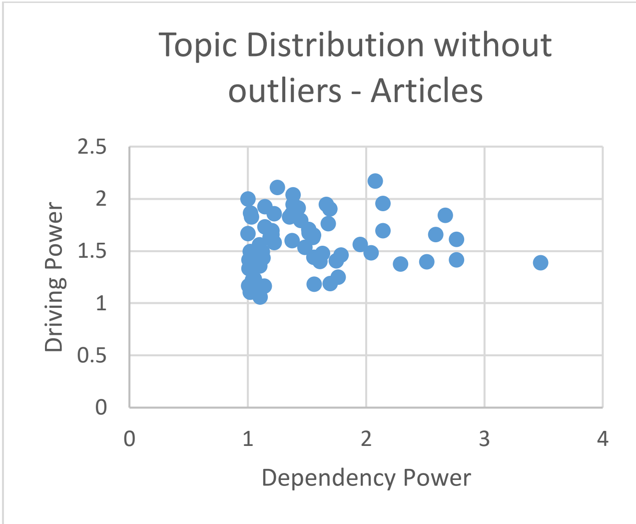

Figure 11: Topic distribution of articles without outliers

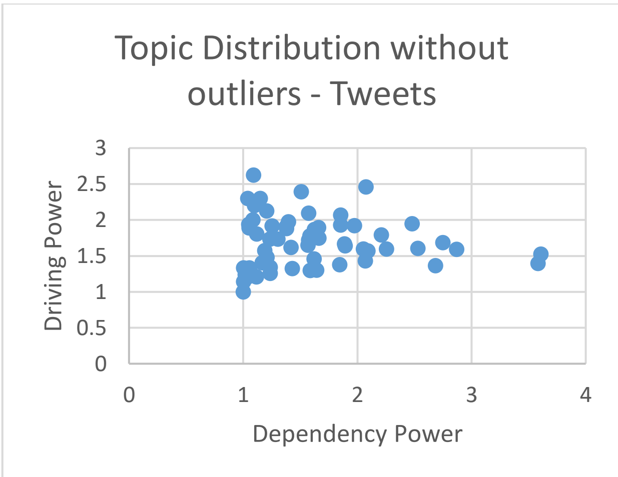

Figure 12: Topic distribution of Tweets without outliers

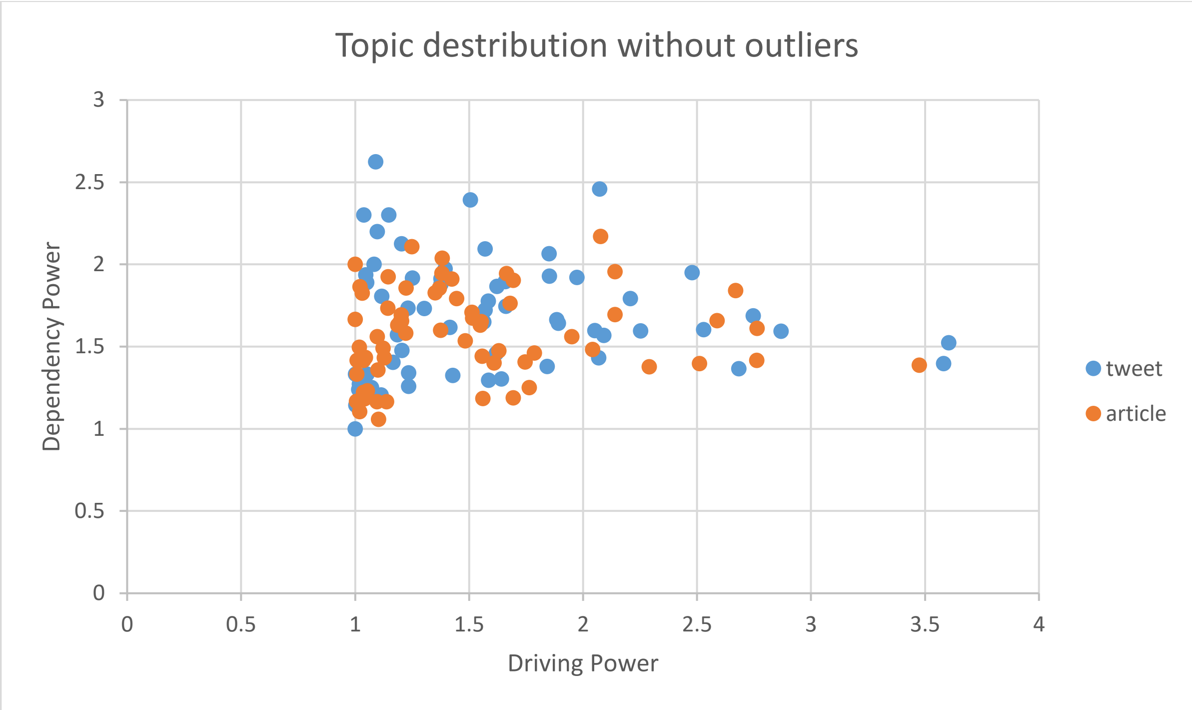

Figure 13: Topic distribution without outliers- Twitter and Articles

73 topics which have non-zero value are categorized into 13 categories including "Abroad"," Background"," Countries", "Demographics", "Economic", "Financial", "Globalization", "Host", "Job", "Network", "Origin", and "Political". Table 3 shows all the topics and the categories they assigned.

Table 3: Topics and assigned categories

No.	Topic	Category	No.	Topic	Category
1	abroad	Abroad	**38**	enrollment	Education
2	foreigner	Abroad	**39**	knowledge	Education
3	International	Abroad	**40**	school	Education
4	overseas	Abroad	**41**	student	Education
5	Australia	Country	**42**	university	Education
6	Brazil	Country	**43**	background	Family Background
7	Canada	Country	**44**	family	Family Background
8	China	Country	**45**	parents	Family Background
9	France	Country	**46**	performance	Family Background
10	Germany	Country	**47**	cost	Financial
11	India	Country	**48**	fee	Financial
12	Italy	Country	**49**	growth	Financial
13	Kazakhstan	Country	**50**	cost	Financial
14	Saudi	Country	**51**	fee	Financial
15	Turkey	Country	**52**	growth	Financial
16	UK	Country	**53**	Income	Financial
17	US	Country	**54**	tuition	Financial
18	Vietnam	Country	**55**	global	globalization
19	culture	Demographic	**56**	world	globalization
20	gender	Demographic	**57**	destination	Host
21	language	Demographic	**58**	Europe	Host

22	economic	Economic	**59**	host	Host
23	GDP	Economic	**60**	Job	Job
24	population	Economic	**61**	unemployment	Job
25	poverty	Economic	**62**	work	Job
26	quality	Economic	**63**	link	Network
27	resource	Economic	**64**	media	Network
28	rural	Economic	**65**	Network	Network
29	service	Economic	**67**	relatives	Network
30	urban	Economic	**68**	social	Network
31	utilization	Economic	**69**	home	Origin
32	academic	Education	**70**	Origin	Origin
33	Bologna	Education	**71**	border	Political
34	class	Education	**72**	government	Political
35	degree	Education	**73**	law	Political
36	diploma	Education	**74**	policy	Political
37	education	Education	**75**	security	Political

When we run MICMAC analysis for categorized data, the results can examine easier. **Error! Reference source not found.** presents the categorized data from both Twitter and article datasets. The value of the driving power and dependency power of the categorized variables are very close to each other (The distance between the same topic from two datasets is smaller than 5) and it is observable in **Error! Reference source not found.**. There are two exceptions in this plot which are shown by arrows. Exceptions are two factors that have a considerable distance which means very different driving power and dependency power. The first one is “Family background” and the second is “Education”. Family background in tweets has a Driving power of 20.7 and dependency power of 15.4 while the same factor in the academic article has the driving power of 5.5 and dependency 5.2. The differences between values indicate that people having more words about the family background when they send a post on Twitter compare to the researchers when writing an article about student mobility. In other words, we can say that researchers do not have much concern about the family background in their publications, while general opinion on Twitter puts this factor in a powerful position. Having limited differences in driving and dependency power of the other 11 factors can verify that the results of the Fuzzy- ISM approach for textual data. In **Error! Reference source not found.** we can also compare the power of each category with the others. “Countries” is a linkage variable with the highest driving power and dependency power. “Education” also has high driving power and dependency power which group it into linkage factors. “Economic” and “Abroad” also have high driving power that causes put them into independent factors. The other factors can be classified into Autonomous factors.

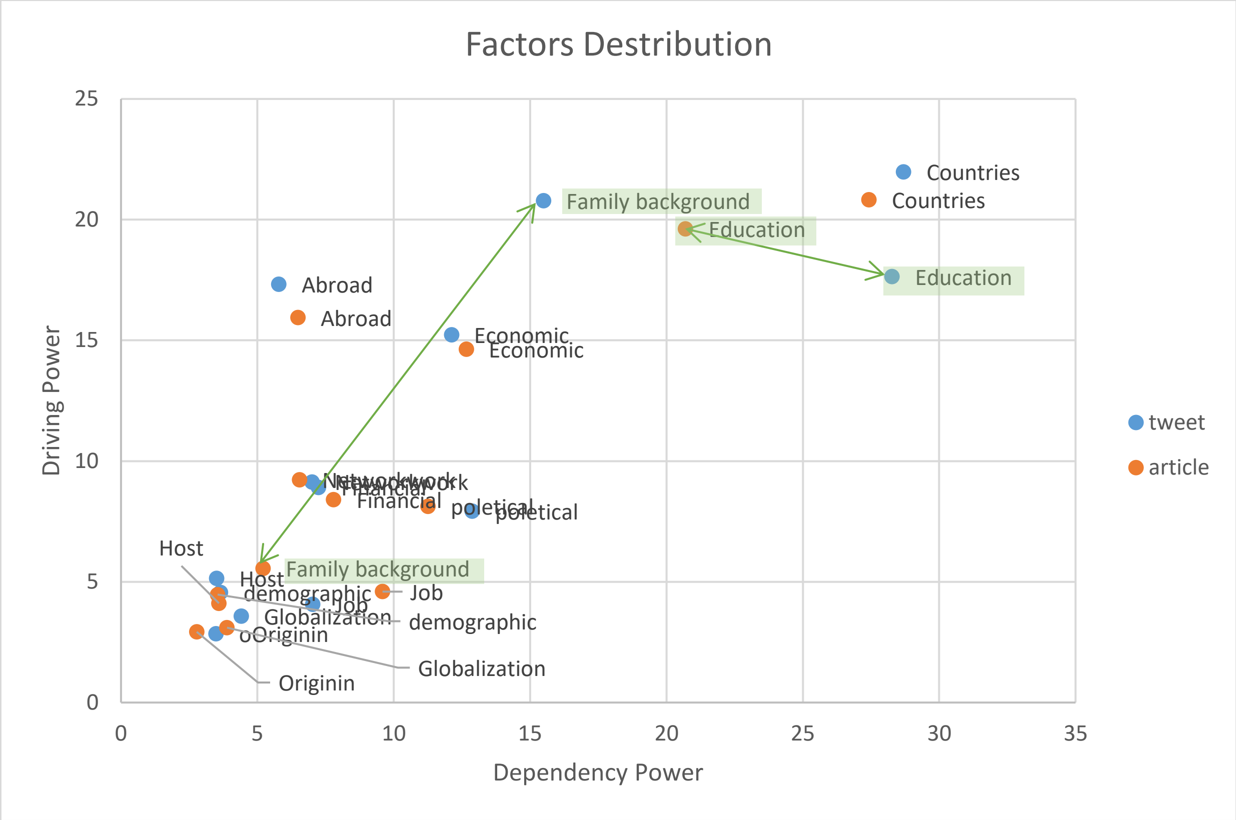

Figure 14: Factor distribution for Twitter and articles

We also run MICMAC analysis for a list of countries in data. Table 4 presents the list of countries that we observed both in academic articles and Twitter. Some of these countries are the common host countries and some of them are famous as the origin countries and there are countries in the list that consider as both host and origin. Therefore, we just identify the driving power and dependency power of each country in Figure 15. As Figure 15 shown, all the 14 countries from tweets and articles have very close power and we cannot see any out range between the two datasets. While the US has a significant dependency and average driving power, it classifies as a linkage country. Germany and China have high driving power, but their dependency power is not significant. The interesting point in Figure 15, there is no variable in the autonomous category. All the countries have driving power between 1.2 and 2 but the range of dependency is varied.

Table 4:List of countries observed in datasets

No.	Country Name	No.	Country Name
1	Australia	8	Italy
2	Brazil	9	Kazakhstan
3	Canada	10	Saudi Arabia
4	China	11	Turkey
5	France	12	The U.K.
6	Germany	13	The U.S.
7	India	14	Vietnam

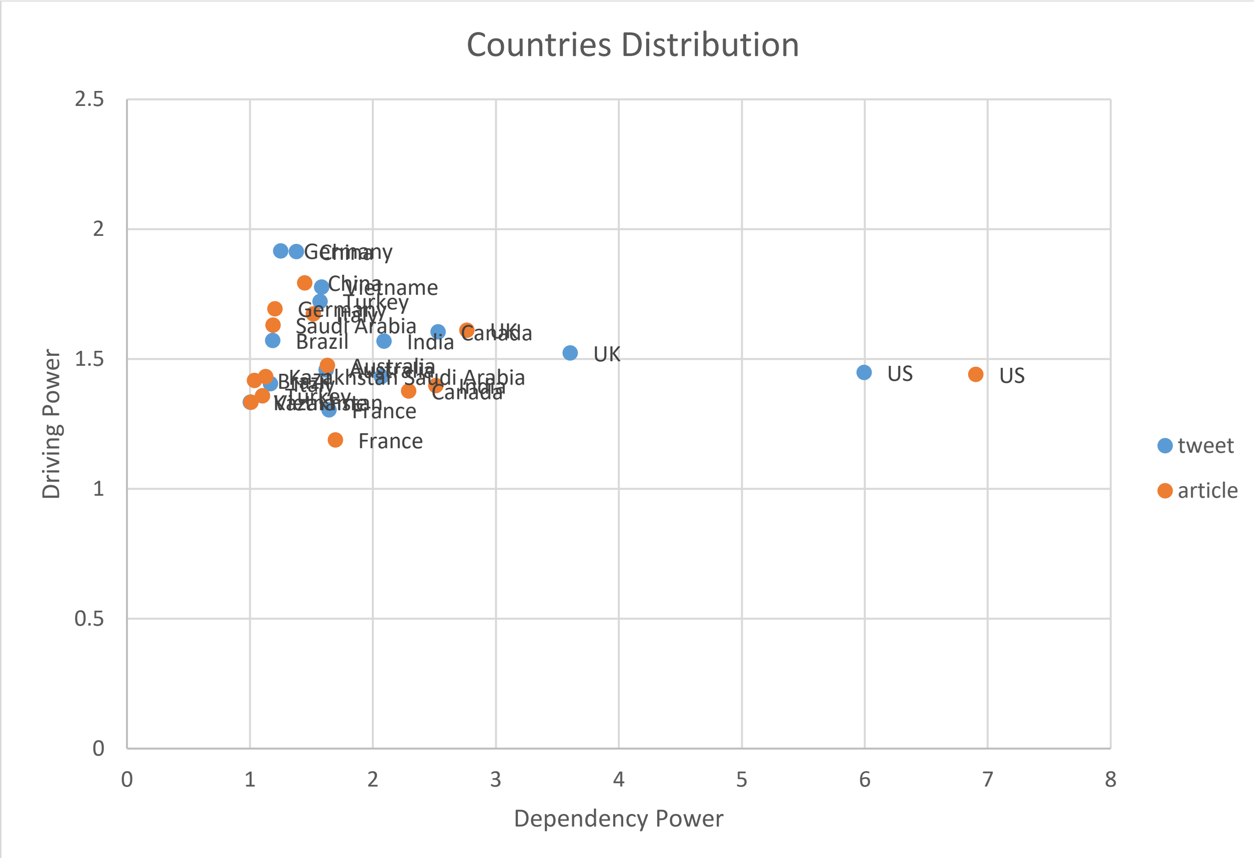

Figure 15: Countries distributions- Twitter and articles

5 Discussions

This section is organized with three main subsections that can cover all aspects of this research. The first is a contribution to the theory that reviews findings that this study adds to the student mobility topic. The second one is a contribution to the methodology which highlights the novelty in the created method. The last part is the implication to the practice which identifies the impact of our work from different aspects.

5.1 Contribution to the Theory:

Researchers from different domains still have doubts about social media trustworthiness as a research data source. The feasibility of using Twitter data for student migration is one of the contributions. Although previously many researchers confirmed the feasibility of social media but using content analysis in this domain is not common. One of the interesting results that should be discussed is the closeness of results from the test data set which is social media data and ground truth which are the articles, such as topic distributions for both data sets that are presented in Figure 13. This figure is built after removing the outliers topics. Considering the outliers also important. There are two common words in outliers of datasets, they are “foreign” and “student”. It is not surprising having these words repeated because the whole topic is about international students or foreign students. The other two outliers of articles are “US” and “work” and another outlier of tweets is “Family”. The differences between these outliers come from a more different point of view of student mobility. In fact, on Twitter people, more talk about the family of international students while researchers may pay less to this topic because of the access to the data or hardness of the study. Removing the outliers shows the distributions of the topic for both datasets that have the same pattern. Although the pattern of distributions cannot show the closeness of the topics one by one, it confirms that both datasets have a very similar shape of distributions. For comparing the distance of topics from two datasets we should look deep at Figure

14. It shows the distance of each set of topics. As shown in Figure 14, almost all topics from one dataset are close to the same topic from another dataset, since two sets of topics have considerable distance. One is “Family background” which has very high driving power and dependency power from tweets in comparison to the data from articles. This also confirmed that people who share posts on social media have more access to their family background and they talk about this data while this type of data for scientists and researchers is hard to capture. Another topic from tweeter which has a distance from its mirror from articles is “Education”. Although the distance between these two is shorter than the distance between “Family background”, it is considerable. “Education” in Tweeter dataset has higher dependency power than “Education” in the article dataset, while their driving power does not have important differences. This fact indicates that people on social media relate education to more numbers of other topics. In simpler words, people talk about more factors while they talk about education. (More factors mean more different types of factors and more repetition of each factor.)

Another observation that should be discussed is the list of countries. The countries from both datasets are presented in Figure 15. There is no considerable distance between the name of each country from datasets. Between these countries, the US has the highest dependency power, which means the US comes after many other factors more than other countries' names. Germany has the highest Driving power and this fact indicates that the name of Germany is followed by many other factors. It can explain two different policies of these two countries to attract immigrant and international students. The US has traditionally attraction for international students because of the famous university and opportunity for having a good job, while in recent years Germany projected a big plan to attract international students.

5.2 Contribution to the Methodology:

The power of the ISM method and its extensions causes the use of this method to expand in a very wide range of domains. However, having a specific domain expert who knows about the ISM method is not easy. The hybrid method (TFISM) proposed in this research enables us to run the ISM method relying on textual data. We extract the effectual factors in each domain and find the potential relationships between factors with the help of text analysis. This technique can apply a huge amount of written documentation to find potential factors and possible relationships. In the original version of ISM experts with limited knowledge or a specific point of view were decided about the factors, while this approach frees us from the limitation by considering vast documents from different perspectives.

5.3 The implication to practice:

Relying on the text makes TFISM approach domain-free. Providing an adequate amount of valid text data enables us to apply TFISM for any domain. We may not have expert knowledge in a specific domain but TFISM with accessing knowledgeable textual sources will help us to have a professional view of the factors and associations and process between them.

Furthermore, the proposed method can highlight the differences between data sources. In other words, TFISM shows the power of similar factors from a different document. This may help for comparing the content for the document with very similar topics. When similar factor grades with different driving power and dependency power, it also can help us to clear the different points of view of content producers.

6 Conclusion

Student mobility has been a matter of very high importance for the government and humanities. In this paper, we propose a very practical method to build a model for student mobility based on textual data. Implementing TFISM as a hybrid methodology enables us to understand the mass text data and extract the relational model between extracted factors. Moreover, this method facilitates the process of analyzing the factors and validating their power that before had been done with the consultations of the domain’s expert. In a comparison of other researches, knowing the power of each factor make a valuable change in the process of decision-making. This method can apply to the textual data from different domains, it also can be applied to many fields of studies and can help us to manage unsupervised textual data. This proposed approach was evaluated by expert knowledge and the multi-level internal process. For the subject of student mobility, the results clarify the power of and influence of each factor.

7 References

[1] Maas, A., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011, June). Learning word vectors for sentiment analysis. In *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies* (pp. 142-150).

[2] Llorente Pinto, A., Garcia-Herranz, M., Cebrian, M., & Moro, E. (2015). Social Media Fingerprints of Unemployment.

[3] Findlay, A. M., Stam, A., King, R., & Ruiz-Gelices, E. (2005). International opportunities: searching for the meaning of student migration. *Geographica Helvetica*, *60*(3), 192-200.

[4] McCallum, A., Wang, X., & Mohanty, N. (2006, June). Joint group and topic discovery from relations and text. In *ICML Workshop on Statistical Network Analysis* (pp. 28-44). Springer, Berlin, Heidelberg.

[5] Ahmad, M., Tang, X. W., Qiu, J. N., & Ahmad, F. (2019). Interpretive Structural Modeling and MICMAC Analysis for Identifying and Benchmarking Significant Factors of Seismic Soil Liquefaction. *Applied Sciences*, *9*(2), 233.

[6] Attri, R., Dev, N., & Sharma, V. (2013). Interpretive structural modeling (ISM) approach: an overview. *Research Journal of Management Sciences*, *2319*, 1171.

[7] Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, *41*(3), 260-271.

[8] Bhandari, R. (2017). A world on the move: Trends in global student mobility.

[9] Castelli, F. (2018). Drivers of migration: why do people move. *Journal of travel medicine*, *25*(1), tay040.

[10] Choudaha, R. (2017). Three waves of international student mobility (1999–2020). *Studies in Higher Education*, *42*(5), 825-832.

[11] Choudaha, R. a. (2012). Choudaha, R. and Kono, Y. New York: World Education Services.

[12] Hughes, C., Zagheni, E., Abel, G. J., Sorichetta, A., Wi'sniowski, A., Weber, I., & Tatem, A. J. (2016). Inferring Migrations: Traditional Methods and New Approaches based on Mobile Phone, Social Media, and other Big Data: Feasibility study on Inferring (labour) mobility and migration in the European Union from big data and social media data.

[13] Karemera, D., Oguledo, V. I., & Davis, B. (2000). A gravity model analysis of international migration to North America. *Applied economics*, *32*(13), 1745-1755.

[14] McGregor, E., & Siegel, M. (2013). *Social media and migration research* (No. 068). United Nations University-Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT).

[15] Simini, F., González, M. C., Maritan, A., & Barabási, A. L. (2012). A universal model for mobility and migration patterns. *Nature*, *484*(7392), 96-100.

[16] Guruz, K. (2011). *Higher education and international student mobility in the global knowledge economy: Revised and updated second edition*. Suny Press.

[17] Hughes, C., Zagheni, E., Abel, G. J., Sorichetta, A., Wi'sniowski, A., Weber, I., & Tatem, A. J. (2016). Inferring Migrations: Traditional Methods and New Approaches based on Mobile Phone, Social Media, and other Big Data: Feasibility study on Inferring (labour) mobility and migration in the European Union from big data and social media data.

[18] Khatwani, G., Singh, S. P., Trivedi, A., & Chauhan, A. (2015). Fuzzy-TISM: A fuzzy extension of TISM for group decision making. *Global Journal of Flexible Systems Management*, *16*(1), 97-112.

[19] Komito, L. (2011). Social media and migration: Virtual community 2.0. *Journal of the American society for information science and technology*, *62*(6), 1075-1086.

[20] Verbik, L., & Lasanowski, V. (2007). International student mobility: Patterns and trends. *World Education News and Reviews*, *20*(10), 1-16.

[21] Hong, L., & Davison, B. D. (2010, July). Empirical study of topic modeling in twitter. In *Proceedings of the first workshop on social media analytics* (pp. 80-88).

[22] Lee, A. H., Kang, H. Y., & Chang, C. C. (2011). An integrated interpretive structural modeling–fuzzy analytic network process–benefits, opportunities, costs and risks model for selecting technologies. *International Journal of Information Technology & Decision Making*, *10*(05), 843-871.

[23] Beine, M., Noël, R., & Ragot, L. (2014). Determinants of the international mobility of students. *Economics of Education review*, *41*, 40-54.

[24] Migration, I. O. (2019). WORLD MIGRATION REPORT 2020. Geneva: International Organization for Migration.

[25] MSc., D. M. (2000). Interpretive Structural Modeling (ISM) . White paper.

[26] Grover, P., & Kar, A. K. (2020). User engagement for mobile payment service providers–introducing the social media engagement model. *Journal of Retailing and Consumer Services*, *53*.

[27] Razavisousan, R., & Joshi, K. P. (2019, December). Two Tier Analysis of Social Media Collaboration for Student Migration. In *2019 IEEE 5th International Conference on Collaboration and Internet Computing (CIC)* (pp. 158-167). IEEE.

[28] Sandbhor, S., & Botre, R. (2014). Applying total interpretive structural modeling to study factors affecting construction labour productivity. *Australasian Journal of Construction Economics and Building, The*, *14*(1), 20.

[29] Shrimali, A. K. (2019). Interpretive structural modeling of identified success factors to lean implementation in SMEs. *Quality Production Improvement-QPI*, *1*(1), 69-77.

[30] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

[31] Tobler, W. (1995). Migration: Ravenstein, thornthwaite, and beyond. *Urban Geography*, *16*(4), 327-343.

[32] United Nations, D. o. (2017). International Migration Report 2017. New York: UnitedNations.

[33] Valmohammadi, C., & Dashti, S. (2016). Using interpretive structural modeling and fuzzy analytical process to identify and prioritize the interactive barriers of e-commerce implementation. Information & Management, 53(2), 157-168.

[34] Warfield, J. (1979). " History and Applications of Interpretive Structural Modeling," May 23, 1979.

[35] Warfield, J. N., & Cárdenas, A. R. (1994). *A handbook of interactive management* (p. 338). Ames: Iowa State University Press.

[36] Zadeh, L. A., Klir, G. J., & Yuan, B. (1996). *Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers* (Vol. 6). World Scientific.