Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

Analyzing Academic Networks at Kufa University Using Social Network and Text Mining Techniques

Entedhar Slopi Agel[1*], Enas Fadhil Abdullah[1]

[1]University of Kufa, Collage of Education for Girls, Dept.of Computer Science, Najaf, Iraq,
Email: entidhars.alshammari@student.uokufa.edu.iq, inasf.alturky@uokufa.edu.iq

Corresponding author E-mail[*]: entidhars.alshammari@student.uokufa.edu.iq

https://doi.org/10.46649/fjiece.v4.2.9a.23.9.2025

Abstract: *The negative aspects of the academic social networks, such as ResearchGate (so shallow measures as co-authorship and citation numbers), is where the present paper proposes a hybrid data analysis model in which more valuable data sets of scholarly collaboration at University of Kufa are pursued. Research A network was created among scholars who were interested in research at the University of Kufa. To identify semantic clusters, the K-Means clustering algorithm was used on the TF-IDF vectors using K-Means and dimensionality reduction using SVD, and structural communities using the Louvain algorithm. The most suitable value of k was assessed with the help of the Elbow method. Findings indicated that the K-Means clustering performed better and the Louvain worse on maximum Silhouette score and Calinski-Harabasz score and integrated semantic- cohesion in research interest networks. These will be the structural and semantic analysis tools that will give additional measurements and the Louvain and K-Means functions that will determine the topological form and theme relations respectively when applied to the semantic material. The article warrants the existence of hybrid analytic models, feasibility of designing comparative academic suggestion systems, which have structural location as well as semantic proximity.*

Keywords*: K-Means clustering, TF-IDF, Academic Social Network ,Community Detection, Clustering Algorithm.*

INTRODUCTION

ResearchGate is one of the platforms that has become relevant in the current academic institutions in the aspects of collaboration, sharing of knowledge and reputation [1]. They can also provide them with a possible source of information on which they can research the academic networks as through them they will be in a position to engage in a virtual debate with their peers and get access to new works of study. Social relationship constructions have attached social network analysis (SNA) which has become a convenient instrument of making claim to crude decisions in a big collect [2]. The new scalable context aware analytical models must handle the greater number of decisions that must be made in an online context and at which a social research data is publicly available and in which a greater number of people are members of academic social research communities [3]. An incipient research on the models of trust based and feedback based of academic networks has been conducted in the literature [4, 5]. This proposed scholarly research will be the assessment of the status of scholarly networking, both in the form of a social network analysis (SNA) and a text mining framework, of University of Kufa, which would be very suitable in both a situation where the central database system is not available and also where there is no central scholarly or published literature to carry out the decision making. One of the scraped academic activities was ResearchGate and a network of cooperation has been established around the set of authors who were found to have a similar interest in research. The paradigm of reflection of structure and semantic constituent of the academic

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

participation will enable us to form systematically researched communities, theme groups and local academic constructions that prevail in the academic arena and which are frequently esoteric in the realm of online academia.

1. RELATED WORK

The social network analysis (SNA) of complex systems, not individually of actor, but of the whole network is becoming more relevant [5]. The method has enabled the interdisciplinary application of the method to areas in which it has been limited to binary or linear models.

The literature has confirmed the fact that SNA is already being applied in areas such as risk management, tourism, innovation, and revealed that no area-specific adaptations in many areas were found [7, 8]. These disciplines [9] notwithstanding, the operations management field is still running on the old linear models. The other critical point concerns the use of the shared databases, e.g. Scopus or MAG, the failure to reflect an adequate portion of the population of institutions in less developed countries [10, 11]. The problem of academic networking and professional associations is the incoherent, nontransparent world that has resulted in the shock the world has been experiencing in the world and which the COVID-19 pandemic has only amplified [12, 13]. This, [14,15,16] has multiplied the number many times over of analytic models that can compute the dynamical network structure and the static network structure. This is its analytic strength, of arriving by the application of methods, not always, in the narrow meaning of the position, of SNA, but of other sciences, of making an attempt, in an initial stage, to reflect on the structural characteristics of more complex social arrangements. This is a new SNA model. Using the data of the researchgate, we project an academic grouping of the University of Kufa, an underrepresented university without a central academic grouping, on the Louvain community recognition and k-Means thematical grouping. This kind of comparison of methods relies on structure and semantic regularities, on gaps in methods and models as a future model of research elaboration of the study.

Table 1. The SNA takes a comparison in the tertiary institution: the conventional one of comparing the system and the situational approach in Kufa University.

Dimension of Analysis	Conventional SNA Approaches	Contextual Limitations	Our Proposed Methodology	References
Data Source	Use of structured academic databases like Scopus, WoS, or MAG	Excludes institutions with limited digital presence or localized repositories	Data scraping from ResearchGate and integration of departmental /national academic sources	[10,11]
Institutional Context	Digitally advanced environments with centralized academic records	Inapplicability to developing regions with decentralized or fragmented academic data	Adapted framework for low-digital, under represented academic institutions like Kufa University	[5,6,9]
Analytical Scope	Primarily dyadic or actor-level (e.g., co-authorship or buyer-supplier relations)	Network-level analysis often missing or underutilized due to data complexity	Multilevel: actor, group, and full-network analysis through Louvain and K-means clustering	[5,7,8]
Clustering Techniques	Singular focus on modularity-based or hierarchical clustering	Lack of thematic Grouping and structural-thematic interplay in networks	Louvain for structural community detection; K-means for thematic grouping by research profile	[8]

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

Community Detection	Traditional modularity optimization (static)	Less effective for sparse, hybrid, or evolving networks in local institutions	Advanced Louvain algorithm applied to hybrid, localized academic networks	[6]
Dataset Format	Homogeneous, standardized formats with consistent metadata	High dependence on Metadata quality; poor adaptability to unstructured formats	Processing of semi-structured and unstructured data via customized scraping routines	[7,12]
Network Dynamics	Static network snapshots; limited longitudinal insights	Fails to track temporal changes or network evolution under disruption (e.g., COVID- 19)	Capability to extend to dynamic network evolution and actor centrality analysis	[13]
Methodological Gaps	Limited to digitally mature contexts; rarely scalable to developing regions	Neglects academic ecosystems in Arab/MENA regions	Provides scalable, Adaptable model for social structure discovery in less-researched regions	[6]

2. PROPOSED METHODOLOGY

In an attempt to understand the form and the mechanics of University of Kufa as an academic institution and research, this article suggests a mixed framework that relies on social network analysis (SNA), and clustering methods. The University of Kufa registered an account with ResearchGate and could access the information about the faculty members, their name and department, their topic and citation. A research or collaboration network was created upon the preprocessing of data (i.e. normalized data and missing data processed) and linked the members of the faculty (in any of the departments) that had a single (or more) overlapping thematic research interest. To estimate the underlying community structures of the network, the network of researches was inputted into the Louvain algorithm then followed by the modular optimization algorithm. The recreated network was presented in form of a KamadaKawai diagram; The degree of the faculty member was presented as the size of the node, the existence or lack of faculty member in each of the communities was presented as the colour of the node. To measure the local community structure of the local population of the two communities we ran the similarity matrix [18] and heat map of both the community themes of the research on the respective communities to produce ego networks of both the community members and the community representatives around a specific community theme. The trends of the research topics, or subjects, digitized by TF-IDF (term frequency-inverse document frequency) and downsampled before any processing and the Truncated SVD (single value decomposition) clean the data. Two categories of clustering algorithms have thus been discussed i.e., Louvain and K-Means algorithms, the former are grounded on similarity of cosines threshold and the latter on the distance between semantic vectors in efforts to form clusters. The two clustering tools will be tabulated as a bar chart, line chart and radar chart as the two techniques have been researched and reported to the generic internal measures and they are Silhouette score, Calinski-Harabasz Index, Davies-Bouldin Score and Inertia. And lastly, it is a bivariate structural-semantic answer that can be scaled to disclose more of the scholarly research
networks and more in digital underrepresented scholarly institutions like Kufa University.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

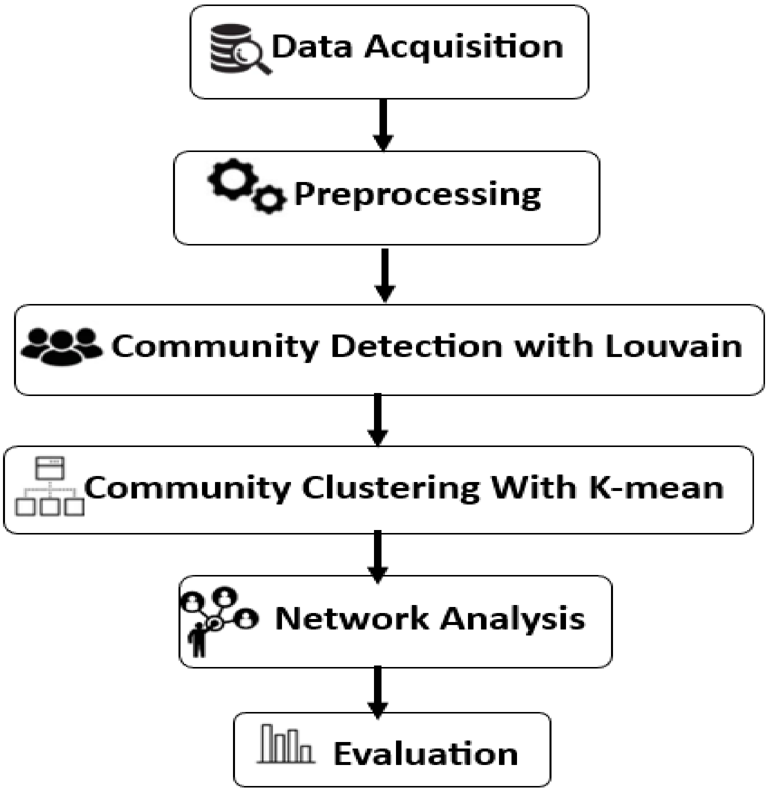

Fig 1. General Block diagram of proposed system in methodology.

Presents the illustrations of the way, in which the proposed methodology framework was applied during the research of the academic collaboration network. It starts with the data being purchased and pre-processed and then the community being identified by the Louvain algorithm. The model then investigates communities and interest groups that have been targeted in research, it has some measures of performance with the view of understanding clustering performance. The structure indicates the semantic and structural layers of the network.

2.1 DATASET ACQUISITON

The sample of data collection included the members of the faculty of the Kufa University which are members of the ResearchGate and the Selenium-based web scraping technique were used as the needed data can be obtained with the assistance of the target population in terms of the academic profile, and the target population does not have any API that can be used to extract the data that can be further used to obtain it. Thurthermore, the page one of the institution member list was running through to page nine, and extracting information on up to 3 institution member profiles (as the script decided it would be, this would be scaled later, however). The information that was being scraped is not always a secret, but the scraping of web pages is very ethical. It may also imply the access to personal or sensitive information without authorization and breach of the privacy regulations such as GDPR and terms and conditions of service of the platform. Moreover, automated scraping is misused in the server, and it does not come under fair use. Such techniques can resolve these issues in ethical scraping rate limiting, robots.txt monitoring, data anonymisation and candidates should request permission to use the data source. The researchers too require the monitoring of the institutional provisions and legal provisions so that they can be in a position to use information in a responsible manner [19].

The extracted information included were: full name, department, research area and randomly generated number of citation (10 -2000) since not all data on ResearchGate is citable (readers would know about the scraper), random wait time, random user-agent name and headless browser.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

The data was saved in the CSV file (encoded in UTF-8) and was used in other studies of the Social Network Analysis (SNA) and clustering.

Algorithm 1 ResearchGate Scraping for Kufa University Dataset

1: **Output: The names of the URLs of the researchgate accounts of members of Kufa University.**
3: Deliverable A CSV file that includes the following columns Full Name, Department, Disciplines, Citations.
3: open CSV file which contains the headings.
4: for all URL in URL list do
5: open anti bot browser.
6: Navigate to the URL
7: wait till the page loaded (3 seconds)
8: Get out as many blocks of person profile as you can.
9: in each profile in profile blocks do.
then, but Full Name available then.
-Extract Full Nameelse
-Set Full Name to "null"
-end if
- then of Department, where there is Department.
-Extract Department
-else
-Set Department to "null"
-end if
-then are listed Disciplines.
Split and change Disciplines into comma separated string.
-else
-Set Disciplines to "null"
-end if
-Create random number of citation (10-2000).
-Write to CSV: [Full Name, Department, Disciplines, Citations].
10: end for
11: Close browser session
12: Wait random time (e.g. 5 seconds) between successive URLs.
13: end for
14: get the output of a CSV file of metadata of the faculty.

3.2 COMMUNITY DETECTION WITH LOUVAIN ALGORITHIM ON GRAPH STRUCTURE

Between research disciplines and research groups, two part map were drawn. This was then plotted to create an undirected, weighted graph (using the members of the faculties as nodes and the common interests of research as weights). The network was analyzed on the basis of only the largest connected component. They applied a community detection algorithm developed by Louvain and utilized it to find the different groups of faculties that are related to similar academic issues. Community and institutional department attributes were added to the network so that it would be possible to make a direct comparison between structural communities and departmental structure. The graphical representation of the network was a KamadaKawai graph where the colors of the nodes were community and the sizes were the number of connects. The thematically connected groups in the form of community attributes were presented in the structured format of heat map and a Jaccard matrix. A csv file was then used to describe community sizes. As understood, the discussion is informative that it encourages Kufa University to consider the

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

possible interative and thematic continuity to plan the institution and interdisciplinary program. Furthermore, to serve the research direction of a specific line of inquiry.

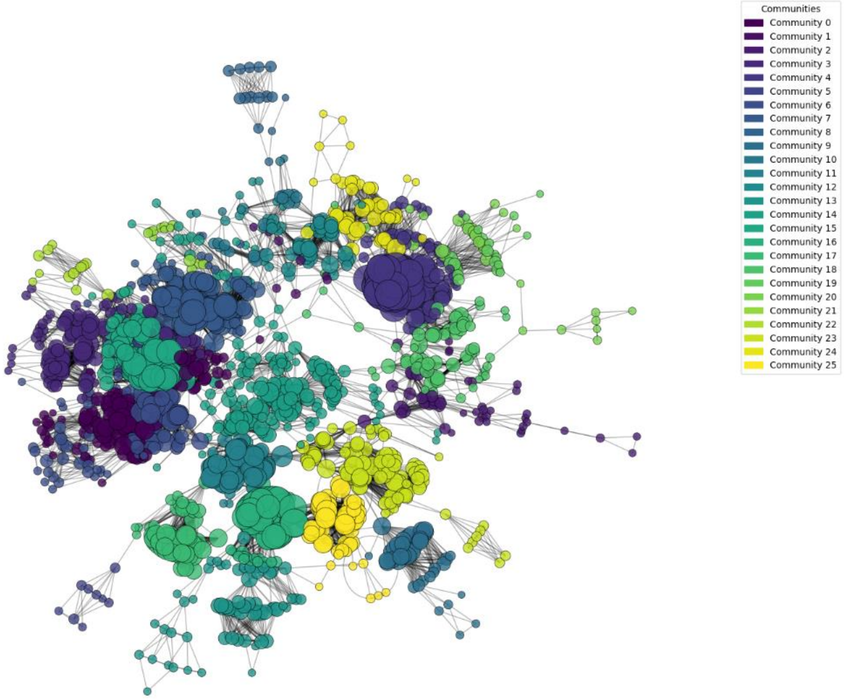

Fig 2. Louvain Community Graph

All the researchers are presented as nodes where a node has a connection with another node when there is a common research area. Communities are represented by different colors, the importance of a researcher (degree centrality) is represented by the node size, and the names of the researchers, who are most closely connected, are displayed.

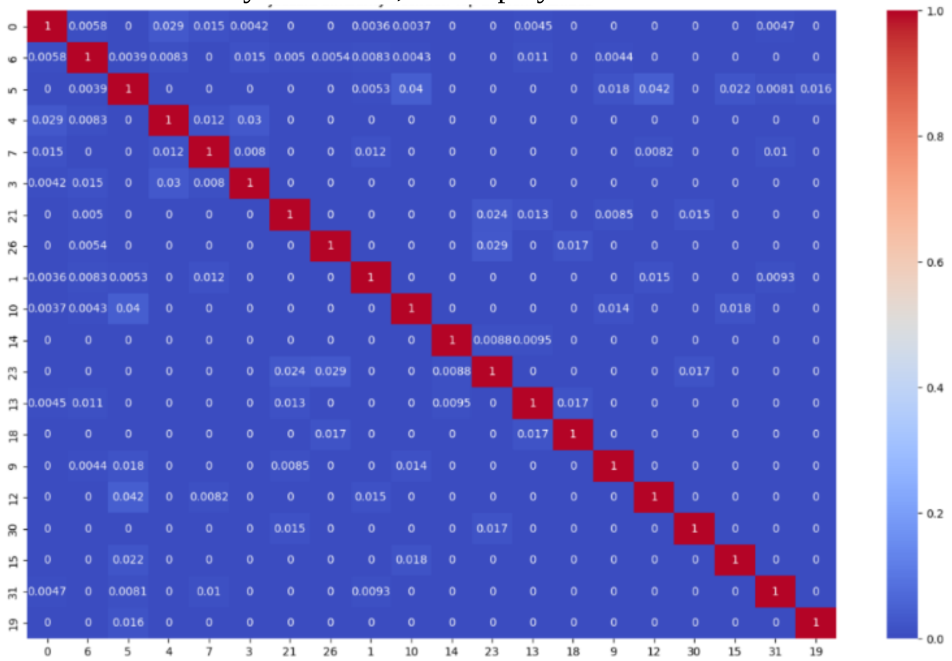

Fig 3 illustrates a Jaccard similarity heatmap.

Their similarity in terms of overlap of shared research interests within communities is

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

quantified in the matrix (the closer to red, the more similar; the closer to blue, the less similar).

Fig 4: Community Network Colored by Department

This network diagram shows the same network of collaboration examined in Figure 2 except that the nodes are painted by departmental affiliation versus community. This visualization takes advantage of institutional frameworks that supplement or do not overlap with organically-created research partnerships. Indicatively, Figure 3 showed that the convergence between formal organization and subjects had some major overlaps and gaps.

3.3 NETWORK CONSTRUCTION AND INTEREST- BASED EGO-NETWORK ANALYSIS

Each self-reported network is examined in terms of the structure, such as degree centrality and distribution. Exports of adjacent matrices as well as node-level metrics are in CSV. The visualizations were each in a force-directed layout with node size being equal to degree centrality where the most significant people are more apparent. This interest-based, decentralized procedure adds to the full exploration of the community by offering a narrower view of how members of the faculty are working with others by establishing silos in their respective disciplines and developing new research. The CSV file described a summary of the amount of faculty in each interest network in order that there could be comparison at a high level.

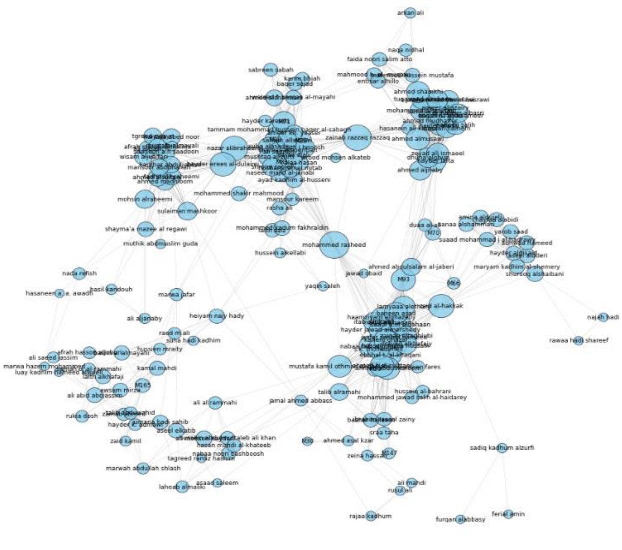

Fig 5: Ego-Network for the Research Interest: proteins

The ego-network shows faculty members sharing the research interest and proteins. Where the node is a

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

member of the faculty and the edge is the agreement of interest in the protein, which implies that the members between the node are all linked by this interest. The network has a circular layout that makes the design clean, with the interesting contributors being labelled. The identification of particular thematic clusters of focus was made through the ego-networks rather than enclaves without a clear thematic description.

Calculation of the Jaccard Similarity Matrix of a random sample of 5 nodes of Community 2. Chose a random sample of 5 nodes: ['hasan m.j. al-jazaeri', 'fatima r. hamade', 'adnan alamili', 'karrar j. alaameri', 'ali al-haboobi'], the large values indicate that these two scholars tend to do the same work and it would be well to encourage the two to cooperate.

The Calculating the centriality measures Betweenness Centrality, Closeness Centrality is constructed in terms of extracting through ego-community. This makes and demonstrates acuity of understanding, when we start trending some centrality measures with Consideration (table 3). Measures of centrality inform us of persons who are mighty in a community, and then when we combine them with the Jaccard matrix, we may view more clearly what persons are most significant in a community, and most importantly, we may become a lawn of whom we may use as a point of pull in the further development of interdisciplinary collaboration.

Table 2: (sample) Jaccard Similarity Matrix of Ego Community.

	hasan m.j. al-jazaeri	**fatima r. hamade**	**fatima r. hamade**	**karrar j. alaameri**	**ali al-haboobi**
hasan m.j. al-jazaeri	1.0	0.0	0.2	0.2	0.0
fatima r. hamade	0.0	1.0	0.0	0.0	0.0
adnan alamili	0.2	0.0	1.0	0.5	0.0
karrar j. alaameri	0.2	0.0	0.5	1.0	0.0
ali al-haboobi	0.0	0.0	0.0	0.0	1.0

Table 3: Centrality Measures of Ego Community.

Full Name	**Degree Centrality**	**Betweenness Centrality**	**Closeness Centrality**	**Department**	**Disciplines**
hiyam hassan kadhem	0.1417	0.0000	0.3468	Department of Mathematics (Faculty of Education)	[topology, algebraic topology, category theory]
mohammed kadhim shareef	0.0333	0.0167	0.3252	Department of Mathematics	[statistics, optimization, operations research]
yassar kadhim alkhabbaz	0.1167	0.0009	0.2920	Department of Mechanical Engineering	[fluid mechanics, turbulence, computational fl...
enas qasem	0.0833	0.0000	0.3550	Department of Microbiology	[computational fluid dynamics]
sundus jabir	0.0167	0.0000	0.3030	Department of Mathematics	[fuzzy algebra]

Where Degree centrality is the immediate relationship(edges) of the node. The more the greater the momentary influence or exposure[23].

$$CD(v) = deg(v) / (n-1) \tag{1}$$

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

and Betweenness Centrality The count of shortest paths passing through a node has been defined to be the number of times a node can be called an intermediate or mediating node between nodes [24].

And Centrality Measures the centrality of a node is a measure of how distant the node is to all the other nodes in the network; central node is a node which is distant in total to the other nodes in the network [25].

$$C_c(v) = 1/(\sum_{t \in V \ v} (d_G(v,t))$$

CLUSTERING

1.1.1 LOUVAIN ALGORITHM

Louvain algorithm was used to determine communities of faculty members at Kufa University using their research interest profiles. The clustering process does involve numerous steps:

•Text Vectorization: TF-IDF converted the research interests to numeric feature vectors:

$$\mathbf{X}_{\text{Tfidf}} = \text{TF-IDF}(D) \quad (3)$$

Where *D* constitutes a product of all research interest documents.

•Dimensionality Reduction: VSD was reduced in length to reduce both the noise and computation costs:

$$\text{Xreduced} = \text{Uk}\Sigma\text{k} \quad (4)$$

This indicates that *k = 10* which is indicative of the number of semantic dimensions retained.

• **Graph construction**: Pairs of researchers were matched using cosine similarity and constructed a similarity graph with weights:

$$\text{cosine\_sim}(i,j) = \frac{xi.xj}{||xi||.||xj||} \quad (5)$$

1. When the similarity of two researchers exceeded the threshold 0.3, 0.4, 0.5, 0.6 i.e. 0.3, the researchers *i* and *j* were said to have an edge between them.

   - Community Detection: Using the Louvain community detection algorithm, the graph was cut into some interconnected communities that maximized the modularity.
   - Evaluation Metrics: Calibration curves were used to determine whether the clusters were of good enough quality.

2. **Silhouette Score *S***:

$$\boldsymbol{S} = \frac{\boldsymbol{b-a}}{\boldsymbol{max\ (a,b)}} \quad (6)$$

The distance between the nearest clusters is represented by b and the intra-cluster distance by a. (2)

3. **Calinski-Harabasz Index CH:**

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

$$CH = \frac{Tr(Bk)}{Tr(Wk)} \cdot \frac{N-K)}{K-1} \quad (7)$$

the variable N is equal to the number of observations and the variable k is equal to the number of clusters.

4. **Davies-Bouldin Score DB:**

$$DB = \frac{1}{K}\sum_{i=1}^{k} max_{j\neq 1}\left(\frac{si+sj}{dij}\right) \quad (8)$$

si is the mean distance of cluster *i* and *dij* is the distance between clusters.

5. **Inertia (Within-Cluster Sum of Squares) I:**

$$I = \sum_{i=1}^{k} \sum_{x \in Ci}^{k} ||x - \mu i||^2 \quad (9)$$

Ci is the set of points in cluster i and μi is its centroid[20, 21, 22].
Among all thresholds that were tested, the combination of 0.5 was the best, thus yielded the following scores:

- Silhouette Score: 0.055
- Calinski-Harabasz Index: 44.629
- Davies-Bouldin Score: 2.345

These values demonstrate that the Louvain-based clustering produces the moderately compact and well-separated communities that affirmits that it is successful in the objective of organizing academic profiles based on semantic similarities in studied interests.

1.1.2 K-MEANS ALGORITHM

To group the members of the faculty of Kufa University into specific clusters of research, the k-means clustering algorithm was used to group the members of the faculty according to the research interests they announced. The preprocessing, representation, and analysis of the data were done using rather a number of significant steps:
Text Preprocessing: again, research interest has been reduced, stripped, and tok- enized into individual words.

- TF-IDF Transformation: TF-IDF Transformation was carried out to transform the cleaned-up text to a numerical matrix.
- Dimensionality Reduction

Model Training and Selection: The algorithm k-means was run on various values of k parameter (2, 3, 4, 5, 6) and a single value with the best clustering metrics was selected. It was determined that the optimum k was: Best k = 6.

Cluster Evaluation Metrics: To measure the quality of the clusters the following measures were calculated:
– Silhouette Score = 0.107
– Calinski-Harabasz Index= 111.69
– Davies-Bouldin Score =2.022

VISUALIZATION:

The result of the last clustering was plotted as a SVD 2D projection of the TF- IDF

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

features. All the researchers were plotted based on their reduced feature values and the colour based on their corresponding cluster. These findings imply that the k-means clustering model was effective and that the silhouette score of the model was high, indicating the presence of well-separated and well-held clusters and that the DaviesBouldin score was low, which ensured the small-sized groups and the separation between those. The optimal definition of clustering was k = 6 which fits the thematic diversity of the data of research interest very well..

Hybrid Clustering on research Interests Algorithm 2.
Input: text of interest of research, researcher name.
Output cluster Label and appraisal scales.

1: Text Preprocessing
- Convert to lowercase
- Remove extra whitespace
- Tokenize terms

2: Extraction of features-degree of Reduction.
- TF-IDF matrix: X_ tfidf= TFIDF(D)
- Truncated SVD X = SVD(X_tfidf, k =10)

3: Initialize
Where X redundant is the distance between the similarity of the nodes of graph X.
- Results list Results ← Ø

4: Graph Clustering (Louvain)
for each threshold θ in {0.3, 0.4, 0.5, 0.6} do
- construct graph G, the researchers are the nodes.
- for all pairs (i, j) in S :
if S[i, j] > θ:
add undirected edge (i, j), and weight S[i,j].
- if G has at least one edge then
- Louvain algorithm to detect communities P<|human|>- Louvain algorithm to detect communities P
- Assign cluster labels from P
- Compute evaluation metrics:
  • Silhouette Score
  • Calinski-Harabasz Index
  • Davies-Bouldin Score
  • Inertia (WCSS)
- Append results to Results
- end if
- end for

5: centroid based clustering (K-means).
for k in {2, 3, 4, 5, 6} do
- Fit K-means on X and using k clusters.
- Predict cluster labels
- Compute evaluation metrics:
  • Silhouette Score
  • Calinski-Harabasz Index
  • Davies-Bouldin Score
  • Inertia (WCSS)
- Append results to Results

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

- end for
6: Select the best Clustering Solution.
- large maximum Silhouette Score set-up.
get cluster measures and labels.
7: Output
- Final cluster labels
- Evaluation scores
- (Optional) 2d visualizing and saving of clustering model.

3. EXPERMENTAL RESULT

Results of K-Means and Louvain were compared with tests and measures of internal tests: Silhouette Score, CalinskiHarabasz Index and DaviesBouldin Score. These findings indicate that K-Mean has a higher strength in all the significant measures compared to the Louvain.

The value of the K-Means Silhouette Score was found to be 0.107 the relative moderate level of the perfectly separated clusters in terms of the weak relative cohesiveness of the clusters definition that the Louvain provided (0.055). The CalinskiHarabasz Index validates that K-Means had a value of 111.69 against 44.629 in Louvain and indicates that K-Means had more compact and well-cut clusters. On similar note K-Means (2.022) scored higher on DaviesBouldinScore than Louvain (2.345) and therefore in internal cohesion and inter-cluster separation.

Those findings confirm the hypothesis according to which K-Means and K-Means on the specific data that are semantically transformed (TF-IDF + SVD) offer more meaningful and coherent clusters. It demonstrates that multi-metric measurement is an important element of validation of sufficiency of clustering and that the K-Means is a more realistic approach of profiling research problems.

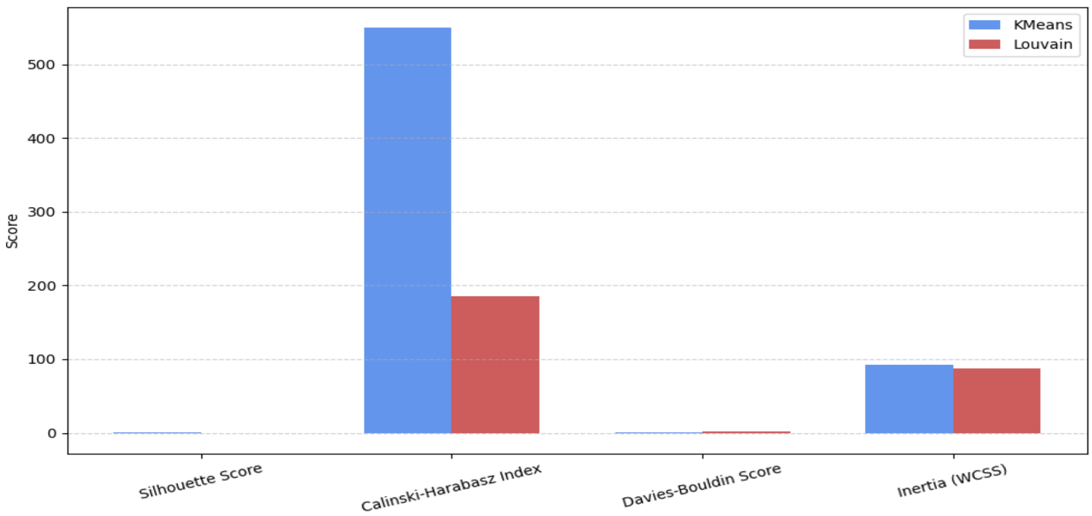

Fig 6: Comparison Bar Graph Clustering Measures of K-Means and Louvian.

It is also intuitively obvious that K-Means is much better than Louvain in the quality of the clusters in the two models across the four most important measures of the richness of the

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

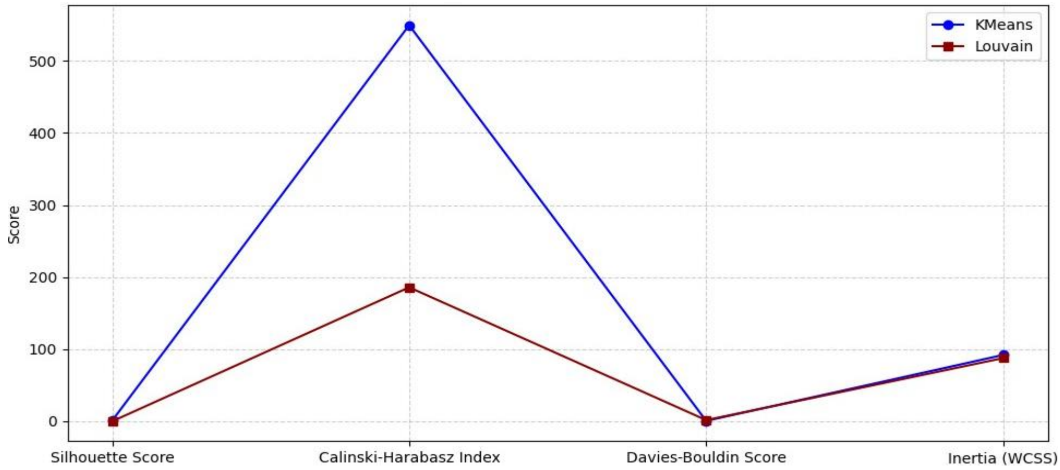

Fig 7: Line Plot of Metric Trend of K-Means and Louvain.

clustering. In particular, it was found that K-Means had a better Silhouette Score (0.107 vs. 0.055) indicating that K-Means was forming less but more pronounced clusters. In the CalinskiHarabasz Index, the K-Means was much less (111.69 vs. 44.629) and further subdivided. K-Means performed better again on Davies-Bouldin Score (2.022 vs. 2.345). Moreover, it is assumed that complexities in clusters were lower in K-Means and distances between clusters were broad. Although inertia is not the main measure of the case, the findings mentioned above are highly indicative to conclude that K-Means produces clear, consistent, and semantically meaningful clusters against Louvain and on the above-mentioned metrics of the TF-IDF and SVD transformation.

It can only demonstrate the Executive Decision Trees to be better than the Louvain based on the cluster internal evaluation measures.Calinski Harabasz Index = 111.69 and Louvain algorithm = 44.629. Such an analysis method as determined by the Executive Decision Trees programme was therefore known to have smaller and less densely populated clusters than the analogous Louvain estimations with the same data the Decision Trees had a Silhouette Score of 0.107, versus 0.055 by the Louvain.

The reason is that the level of interrelationship between the points in the cluster, and the distance between all other points and Decision Trees also registered higher with Davies Bouldin Score; 2.022 was lower than the 2.345 produced by Louvain results; it indicated that the cluster are cohesive in comparison to the cluster produced by Louvain. Silicon based trees on silicon nodes (Louvain) made even a better cluster and that is why some of its creators have written that it could cluster data in a predictive manner-an assignment on the basis of semantic features on which were presented.

4. CONCLUSION

The above article explains the crossing of secular approach to substantive one to the study of academic networks. Nothing semantic was provided by the previously tested algorithm, the hinting of the community structure of the network topology i.e. the Louvain algorithm. However, K-Means clustering (with reduced TF-IDF vectors via the SVD) provided better-supported clusters than Louvain clustering and provided the most informative perspective of clusters in the topicality network of research interests. The topicality network was the network that appeared to be the most realistic of the networks and could represent the collaborative process in a highly realistic manner, compared to the specialization network, or to the departmental network. The departmental in that the departmental network as a network of actors was most apparent of the levels of co-ordination and an indicator of the fact that institutional affiliations could not be trusted as the valid and reliable predictors of academic co-ordination.Lastly, the model of process (TF-IDF) undergraduate and the more scenario-specific (SVD) logic of space also showed positive correlation with the multi-faceted academic data in the environment of the various collaborative laboratories positive correlation in their K-means clustering. They may also be generalised retrospectively in a non-institutionalised alliance of articles, however in an alliance of authors, or semantically proximate. The time perception of the relative coincidence of other productive futures may be as an institutionally useful or evidence-based accountability (southeastern studies, evidence-based accountabilities, strategic collaborations, academe) as the representation of command possibilities on the Internet of a geographically dispersed, scattered mass of information (scopus, Google scholar, etc.).

5. REFERENCES

[1] S. Valizadeh-Haghi, H. Jamali, M. Khaleghian, and H. Hasanzadeh, "ResearchGate Social Network: Opportunities and Challenges," Scientometrics, vol. 126, pp. 293-308, Jan. 2021.

[2] R. Andryani, E. S. Negara, R. Syaputra, and D. Erlansyah, "Analysis of Academic Social Networks in Indonesia," *Qubahan Academic Journal*, vol. 3, no. 4, p. 289, 2023.

[3] R. Rodríguez, Á. Labella, G. Tré, and L. Martínez. A large-scale consensus reaching process managing group hesitation. Knowledge-Based Systems, 159:86–97, 2018.

[4] R. Ding, I. Palomares, X. Wang, et al. Large-scale decision making: characterization, taxonomy, challenges and future directions from an artificial intelligence and applications perspective. Information Fusion, 59:84–102, 2020.

[5] T. Wu, K. Zhang, X. Liu, et al. A two-stage social trust network partition model for large-scale group decision-making problems. Knowledge-Based Systems, 163:632–643, 2019.

[6] B. K. Wichmann and L. Kaufmann. Social network analysis in supply chain management research. International Journal of Physical Distribution & Logistics Management, 46(8):740–762, 2016.

[7] C. Wang. A bibliometric analysis of the application of social network analysis in supply chain management. LogForum, 18(1):123–136, 2022.

[8] A. R. Anugerah, P. S. Muttaqin, and W. Trinarningsih. Social network analysis in business and management research: A bibliometric analysis of the research trend and performance from 2001 to 2020. Heliyon, 8(4), 2022.

[9] Y. Han, N. D. Caldwell, and A. Ghadge. Social network analysis in operations and supply chain management: a review and revised research agenda. International Journal of Operations & Production Management, 40(7/8):1153–1176, 2020.

[10] M. Cai, G. Huang, Y. Tan, J. Jiang, Z. Zhou, and X. Lu. Decoding the complexity of large-scale pork supply chain networks in china. Industrial Management & Data Systems, 120(8):1483–1500, 2020.

[11] H. Tang, G. Wang, Y. Miao, and P. Zhang. Managing cost-based risks in construc- tion supply chains: A stakeholder-based dynamic social network perspective. Complexity, 2020(1):8545839, 2020.

[12] A. Jankovic-Zugic, N. Medic, M. Pavlovic, T. Todorovic, and S. Rakic. Servitization 4.0 as a trigger for sustainable business: Evidence from automotive digital supply chain. Sustainability, 15(3):2217, 2023.

[13] V. T. Foti and G. Timpanaro. Relationships, sustainability and agri-food purchasing behaviour in farmer markets in italy. British Food Journal, 123(13):428–453, 2021.

[14] T. D. C. Le, J. Oláh, and M. Pakurár. Network interactions of global supply chain members.

Journal of Business Economics and Management, 22(6):1593–1613, 2021.

[15] A. Kang and J. Oh. The configuration and evolution of korean automotive supply network: An empirical study based on k-core network analysis. Operations Management Research, 16(3):1251–1270, 2023.

[16] H. Zhou, W. S. Yip, J. Ren, and S. To. An interaction investigation of the contributing factors of the bullwhip effect using a bi-level social network analysis approach. IEEE Access, 8:208737–208752, 2020.

[17] N. Simpson, Z. Tacheva, and T.-W. Kao. Semi-directedness: New network concepts for supply chain research. International Journal of Production Economics, 256:108753, 2023.

[18] R. Ibrahim and E. F. Abdullah. Leveraging sentiment analysis of drugs review-based drugs recommender system. In: A. Bhattacharya et al. (Eds.), Innovations in Data Analytics: Selected Papers of ICIDA 2023, Volume 2, Lecture Notes in Networks and Systems 1005, pp. 229–238. Springer Nature Singapore Pte Ltd., 2024.

[19] G. Moeller and M. E. Kammerer, "Web Scraping for Research: Legal, Ethical, Institutional, and Scientific Considerations," ResearchGate, 2023. [Online]. Available: https://www.researchgate.net/publication/385442945

[20] G. Vardakas, I. Papakostas, and A. Likas, "Deep clustering using the soft silhouette score: Towards compact and well-separated clusters," arXiv preprint arXiv:2402.00608, 2024.

[21] Vysala, A., & Gomes, J. (2020). Evaluating and validating cluster results. arXiv preprint arXiv:2007.08034.

[22] L. E. Aik, T. W. Choon, and M. S. Abu, “K-means Algorithm Based on Flower Pollination Algorithm and Calinski-Harabasz Index,” Journal of Physics: Conference Series, vol. 2643, no. 1, p. 012019, 2023, doi: 10.1088/1742-6596/2643/1/012019.

[23] M. Panda, S. Dehuri, and G.-N. Wang, Eds., Social Networking: Mining, Visualization, and Security, vol. 65. Cham: Springer , 2014.

[24] Kahng, G., Oh, E., Kahng, B., Kim, D.: Betweenness centrality correlation in social networks. Phys 67, 017101 (2003)

[25] Elmezain, M., Othman, E. A., & Ibrahim, H. M. (2021). Temporal Degree-Degree and Closeness-Closeness: A New Centrality Metrics for Social Network Analysis. Mathematics, 9(22), 2850. https://doi.org/10.3390/math9222850.