Article

MOFSocialNet: Exploiting Metal-Organic Framework Relationships via Social Network Analysis

Mehrdad Jalali [ID], Manuel Tsotsalas *[ID] and Christof Wöll *[ID]

Institute of Functional Interfaces (IFG), Karlsruhe Institute of Technology (KIT), Hermann-von Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany; mehrdad.jalali@kit.edu

* Correspondence: manuel.tsotsalas@kit.edu (M.T.); christof.woell@kit.edu (C.W.)

Abstract: The number of metal-organic frameworks (MOF) as well as the number of applications of this material are growing rapidly. With the number of characterized compounds exceeding 100,000, manual sorting becomes impossible. At the same time, the increasing computer power and established use of automated machine learning approaches makes data science tools available, that provide an overview of the MOF chemical space and support the selection of suitable MOFs for a desired application. Among the different data science tools, graph theory approaches, where data generated from numerous real-world applications is represented as a graph (network) of interconnected objects, has been widely used in a variety of scientific fields such as social sciences, health informatics, biological sciences, agricultural sciences and economics. We describe the application of a particular graph theory approach known as social network analysis to MOF materials and highlight the importance of community (group) detection and graph node centrality. In this first application of the social network analysis approach to MOF chemical space, we created MOFSocialNet. This social network is based on the geometrical descriptors of MOFs available in the CoRE-MOFs database. MOFSocialNet can discover communities with similar MOFs structures and identify the most representative MOFs within a given community. In addition, analysis of MOFSocialNet using social network analysis methods can predict MOF properties more accurately than conventional ML tools. The latter advantage is demonstrated for the prediction of gas storage properties, the most important property of these porous reticular networks.

Keywords: metal organic framework; social network analysis; centrality in the graph; community detection

1. Introduction

The modelling and examination of complex systems that contain chemical, biological, ecological, economic, social, technological, and other types of information is a very challenging process if the number of elements in the system becomes very large. Metal-Organic Frameworks (MOFs), a class of chemical compounds composed of metal nodes connected via organic linker molecules, represent a particularly complex example from materials science. The wide variety of metal nodes and organic linker molecules suitable for MOF synthesis and the virtually unlimited number of linker/node combinations lead to an enormous size of the MOF chemical space. While the number of experimentally characterized MOFs already exceeds 100,000 [1], there is no upper limit for the total number of these reticular networks. This diversity in MOF chemistry makes it extremely difficult to navigate through the large design space and to identify MOF materials with suitable properties for a desired application or to identify most representative MOFs for an anticipated study, that cover best the available design space. Identifying an appropriate MOF for a given application in this huge chemical space can, in principle, be carried out by high throughput screening of existing or hypothetical MOFs databases. This screening can be carried our either experimentally or theoretically [2]. However, this approach becomes extremely

costly with increasing size of the database. Recently, new paradigms for the discovery and rational design of materials have been established that are based on Machine Learning (ML) as well as sophisticated data science analysis methods and algorithms [3]. In the field of MOFs, first examples, where ML methods were employed to predict material properties or even to design new MOF structures and predict synthesis conditions have already been successfully demonstrated [4].

The simulation of the adsorption capacity for gases, probably the most important property of MOFs for existing applications, provides an interesting example of ML-based strategies for handing materials classes with large size. The presently most popular ML algorithms are Support Vector Machine, Random Forest and Neural Networks for predicting the absorption of guest molecules by MOFs [5]. In addition, deep learning, a particularly effective ML algorithm, has been used in a number of different applications [6]. For many practical applications, the stability of MOFs in aqueous environments is an important prerequisite and ML-based models could accurately predict water stability of MOFs [7].

When applying ML models, a first important step is the selection of descriptors. The number of the different parameters describing MOF structures is simply too large for a direct, straightforward analysis, therefore also ML-based workflows were introduced, that allow extracting the most valuable descriptors within a given family of MOFs [8]. Combining data mining and machine learning has allowed for the prediction of MOF synthesis [9] and MOF stability [10]. In another study, the authors proposed a machine-learning algorithm to predict the possibility of metal-linker combinations for the guest accessibility of MOFs. In this method, various ML models were evaluated to learn the connection between component chemistry and MOF properties without explicitly requiring a priori knowledge of the MOF structure [11].

In this work, we propose a new approach based on the social network analysis for analysing the chemical space of metal organic frameworks. Social network analysis was originally established in the field of social sciences [12,13], but has been expanded to virtual learning [14], health informatics [15–18], life sciences [19–22], agriculture [23–25], economy [25,26], and others.

Building a social network enables the use of machine learning techniques based on graph mining to extract valuable knowledge from the MOFs data. Our goal in this study is to demonstrate that "social networks" constructed from MOFs are a valuable tool for analysing large MOFs databases, allowing to navigate through the MOF chemical space, identify suitable MOFs for a given application or desired study, and to curate large datasets efficiently.

We used social network analysis (SNA), rather than more traditional machine learning algorithms, since SNA outperforms other ML models in visualizing complex relationships between different MOFs. Additionally, SNA allows to extract information about the properties of a given (e.g., so far unknown) MOF by its relationship to "neighbouring" (known) MOFs in the social graph. SNA therefore allows to extract useful information, e.g., find the most representative MOFs or identify implicit and hidden dependencies between MOFs.

Two primary types of SNA were performed as part of the current research: centrality and community detection. In the centrality analysis, parameters are determined that measure the characteristics of a given MOF node in the graph in relation with other MOF nodes (in this case other MOFs) in the graph. In MOFSocialNet, we deal with different types of centralities, degree centrality, and closeness centrality. Degree centrality focuses on the links of one MOF node to other MOF nodes. MOF nodes with a high degree centrality can be regarded as important MOF structures with similar characteristics to many other MOF structures in the dataset. The closeness centrality is computed by considering the average distance from the target MOF node to the other MOF nodes in the networks. MOF nodes with a high degree centrality can be regarded as very representative MOF structures for the given set of analyzed MOF structures [27]. Node centralities allow to identify the most important or influential node in a graph. For instance, a high value of degree centrality identifies nodes that are in the middle of the network. Thus, by blending the

information provided by the different centralities allows to analyze the MOF networks and to find correlations between MOFs.

One other parameter, which we looked at in MOFSocialNet, is community detection [28–34]. Community detection is essentially a type of clustering problem. Community detection aims to group the nodes according to the relations between them to form strongly related sub-graphs from the entire graph. For example, detection of communities holds an important place in the analysis and functional prediction of the interaction networks between proteins and other molecules in biological cells [21] or to predict and identify disease genes.

In the context of MOFSocialNet, community detection can be applied to provide an overview and a structure to highly diverse MOF datasets. Social network analysis can therefore help to rationalize the categories used to describe MOF types, moving away from "most popular MOF types" towards categories based on more objective features or properties of the MOFs.

2. Methods

In this paper, we construct a social network called MOFSocialNet from geometrical descriptors of MOFs in the CoRE-MOFs database. MOFSocialNet is an undirected, weighted, and heterogeneous social network. Following the construction of MOFSocialNet, we provide a set of social network analytic processes to extract valuable knowledge from the MOF data using graph-mining algorithms. The full workflow of the MOFSocialNet is shown in Figure 1. In the first step, we created a feature vector for each MOF. In the following step, the similarity between each pair of MOF vectors is calculated based on vector similarity methods. In this step we created the MOFs social graph, named MOFSocialNet, where MOFs are the nodes and the similarity between the MOF feature vectors are the links (i.e., the relationships between MOFs). Finally, after removing irrelevant links in the graph, we applied social network analysis methods to extract valuable knowledge from this graph. In subsequent sections, all steps are described in detail.

2.1. Creating MOF Feature Vectors

A feature vector is an n-dimensional vector of numerical features that can e.g., describe an object in pattern recognition using machine learning. In the case of MOFs, every property depends on a set of specific descriptors, i.e., geometrical, chemical, topological, and energy-based descriptors [35]. Before applying the data analysis process, it is critical to identify key descriptors that are highly correlated with the property of interest of the MOF.

For a demonstration of the proposed approach, we limited ourselves to a subset of 1000 MOFs in the CoRE-MOFs database and used eight geometric descriptors for MOF (see Table 1). Therefore, each MOF is assigned to an eight-dimensional feature vector.

Table 1. Geometry Descriptors in CoRE-MOFs Database.

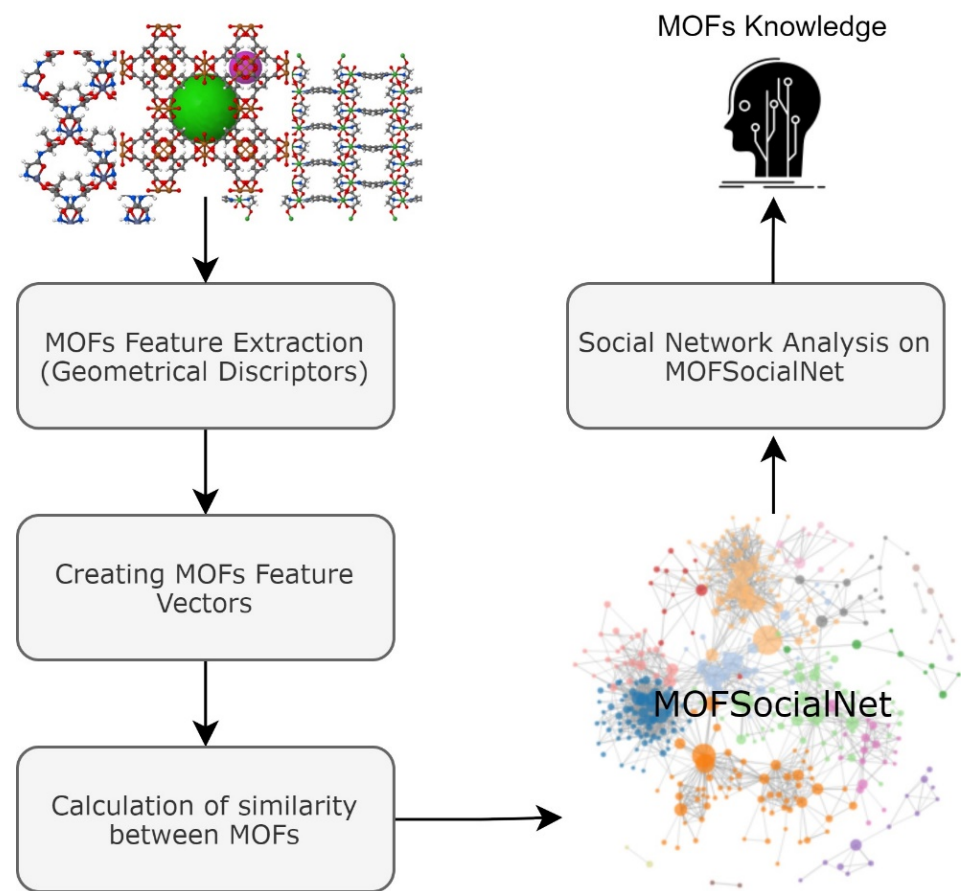| Name of the Descriptors	Symbol
Sphere	Di
Largest free sphere	Df
Largest included sphere along free path	Dif
Crystal density	ρ
Volumetric surface area	VSA
Gravimetric surface area	GSA
Volumetric pore volume	VPOV
Gravimetric pore volume	GPOV

Figure 1. Workflow for the MOFSocialNet creation and analysis.

In order to simplify the further analysis, we normalized the values of the numerical columns in the dataset to a common scale. The most common method of data normalization is Min-Max normalization, which values are transformed into decimals between 0 and 1, using formula 1:

$$v' = \frac{v - min_A}{max_A - min_A} \tag{1}$$

where min_A and max_A denote the minimum and maximum values of the corresponding property A. In the following, the original and normalized values are denoted by v and v', respectively. As a consequence, all v' adopt values from the interval (0,1) [36].

2.2. Constructing MOFSocialNet

We define MOFSocialNet as a graph G = (V,E) [17], where V = {v1, v2, . . . , vn} is the vertex set, and E = {eij = edge from vi to vj | 1 ≤ i, j ≤ n, i ≠ j} is an edge set.

For the construction of the graph, a metric must be introduced to measure the distance between two vertices. In previous works, either direct (Euclidean or Manhattan metrics) or more invoked methods like Pearson's product-moment correlation coefficient (PPMCC) or Cosine methods have been used.

In the current study, we have used the so-called cosine metrics. Given two vectors of MOF descriptors, $\overline{A}$ and $\overline{B}$, the cosine similarity, $\cos(\theta)$, is computed as

$$\text{Similarity} = \cos(\theta) = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}} \tag{2}$$

where A_i and B_i are the components of the feature vectors $\overline{A}$ and $\overline{B}$, respectively. The resulting similarity ranges adopt values from the interval (0,1), with 0 signifying the dissimilarity of the MOFs. A value of 1 reveals that the MOFs A and B are identical.

With this metric, now a graph representing all ~1000 elements of the database can be constructed, consisting of 1000 vertices and 22,000 edges.

In this initial graph, the number of edges is too large for an efficient analysis and an effective strategy to remove weak links must be introduced. Therefore, we removed all edges with a length less than a threshold parameter d.

The representation of the sample network of MOFSocialNet after elimination of weak links using a value for d = 0.9999 (reducing the number of nodes to 2214) is shown in Figure 2. The presence of an edge between two MOFs thus indicates that the similarity is above the threshold value.
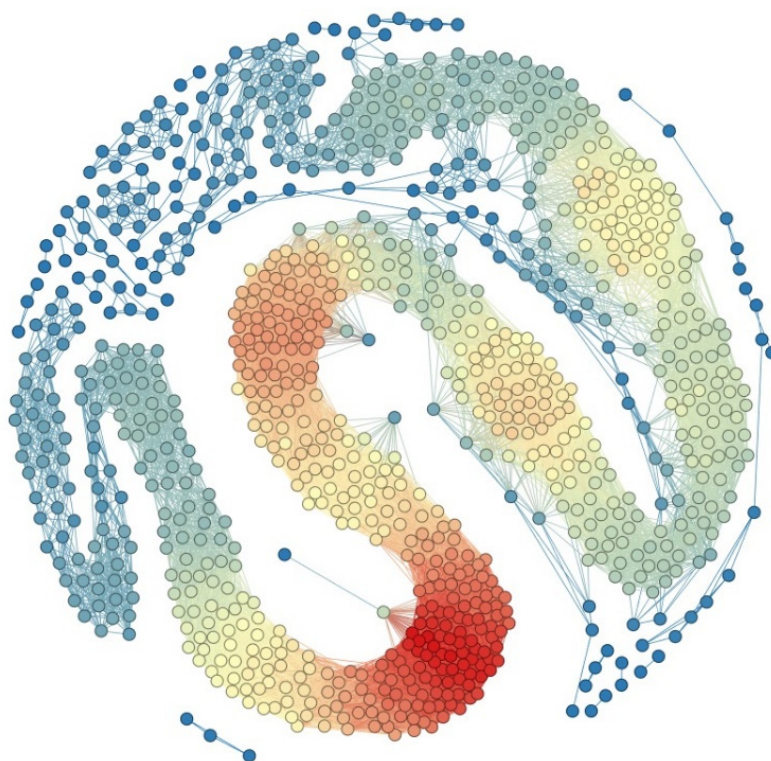

Figure 2. MOFSocialNet with 1000 Metal-Organic Frameworks, the visualization of the MOFSocialNet with Fruchterman Reingold layout in the Gephi software. Different colors indicate MOFs with a similar number of connections to other MOFs.

An important property of a particular node is its degree, the number of edges linking it to other nodes. The probability distribution of these degrees over the whole network is shown in Figure 3. The average degree based on the diagram is 46.

2.3. Centrality Measures in MOFSocialNet

For the next step of the analysis, we determined the centrality of nodes. Centrality means the relative significance of nodes (or vertices) and links (or edges). In our context, centrality measures how similar a MOF is to other MOFs within MOFSocialNet.

The simplest approach is degree centrality, which for a given node is obtained by counting the number of links connecting to other nodes. In MOFSocialNet, a MOF with a high degree centrality has very similar properties to many other MOFs. For the MOFSocialNet graph (G), degree centrality (C_d) is defined as:

$$C_d = \frac{k_i}{n-1} \tag{3}$$

where k_i is the degree (number of edges connected to a node) of node i and n is the total number of nodes in the graph [35].
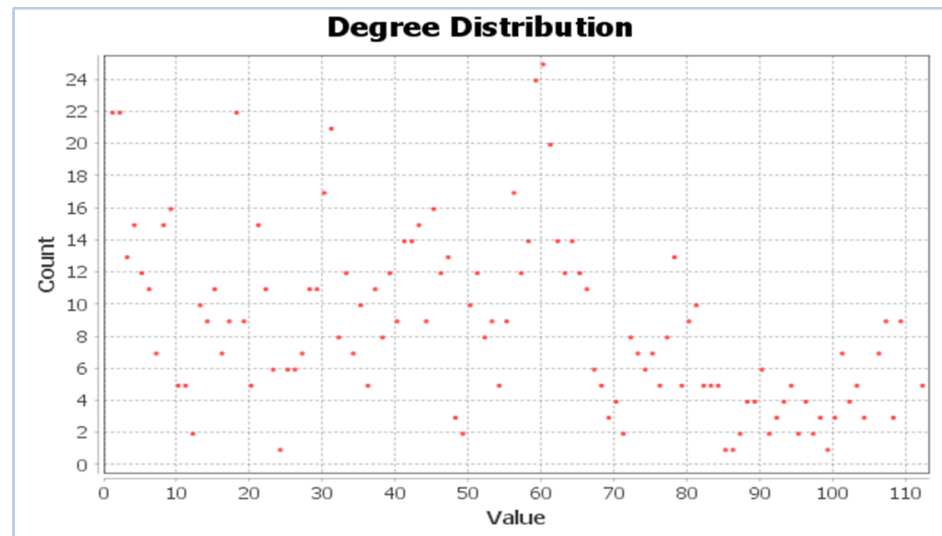

Figure 3. Degree distribution of the MOFSocialNet—the degree distribution is the probability distribution of node degrees over the whole network.

Table 2 presents the 10 MOFs with highest degree and Figure 4 illustrates the MOFSocialNet with these MOFs being highlighted.

Table 2. Ten MOFs with highest degree.

No.	Name of the MOF	Degree
1	LELRUR	112
2	LAZXOB	112
3	KOCWEF	112
4	KOZNIY	112
5	LEYRAK	112
6	j.ica.2016.05.002_2	109
7	JOGHEU	109
8	KEVWUF	109
9	KIJPUQ	109
10	KECRAL10	109

The closeness centrality (C_c) of a vertex is a measure of the closeness of the vertex to the rest of the vertices in a graph. The C_c of a vertex is computed as the inverse of the sum of the hop counts (farness) of the shortest paths from the vertex to the rest of the vertices in the graph. If $d(i, j)$ is the geodesic distance between two vertices v_i and v_j in a graph, then the closeness centrality of a vertex v_i could be computed as the sum of the geodesic distances to the vertices v_j that are in the same component as v_i (a component is the largest set of vertices that are reachable from each other) [37].

Closeness centrality (C_c) is defined as:

$$C_c = \frac{n-1}{\sum_j d(i,j)} \qquad (4)$$

where i is the starting node, j the target node, and $d(i, j)$ is the distance between them. This measures the distance from the starting node to other nodes in the graph [38].

The closeness centrality captures the accessibility of network components. In the MOFSocialNet, being a network of MOF structures, closeness centrality can identify the most representative MOFs of a given dataset. Table 3 and Figure 5 illustrate the network, highlighting the ten MOFs with the highest closeness centrality.
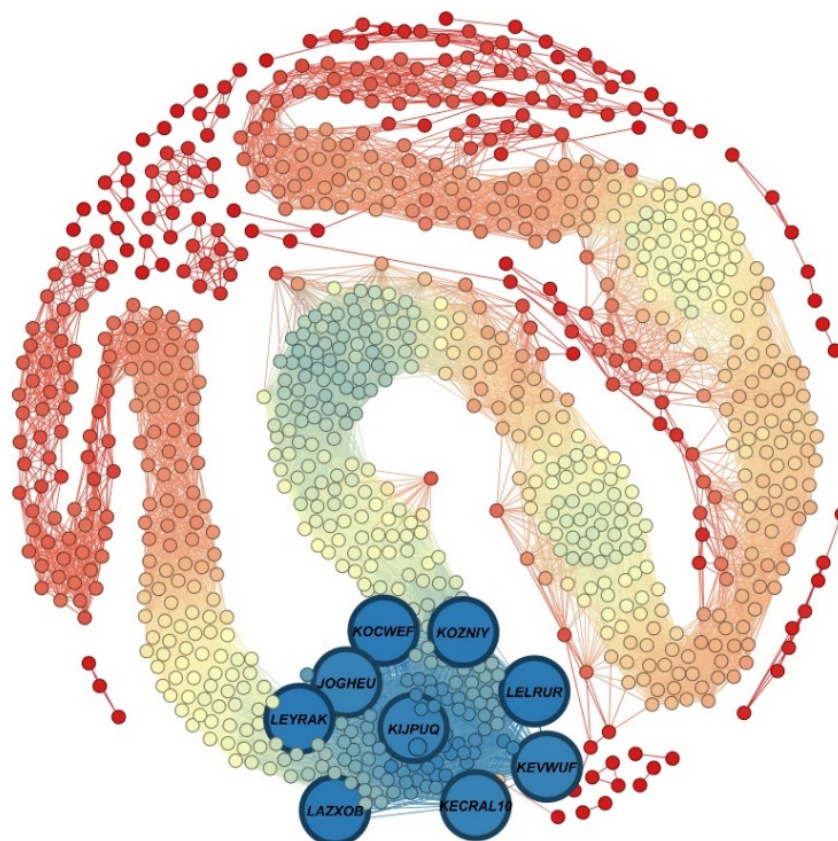

Figure 4. Visualization of MOFSocialNet with the top 10 MOFs with the highest degree, drawn with Fruchterman Reingold layout in the Gephi software. The blue nodes in bold indicate the MOFs with the highest degree.

Table 3. Ten MOFs with highest closeness centrality.

No.	Name of the MOF	Closeness
1	KOSLIP	0.122975
2	KUFVIS	0.122958
3	LARPEC	0.122584
4	KOSLUB	0.122398
5	KOSLOV	0.122331
6	LOSYAV	0.122213
7	KAVTEI	0.122179
8	LEJRIC	0.122179
9	KUNRAO	0.122163
10	AWAKEQ	0.122028

2.4. Community Detection in MOFSocialNet

At the most abstract level, given a Social network G = (V,E), a community can be defined as a subgraph of the network including a set VC ⊆ V of Social Network entities that are associated with a common element of interest. This element can be a topic, a person of the real world, a place, an event, an activity, or a material such as metal organic framework.

Community detection is a common method in the social graph to categorize a large graph in sub-group with similar features or properties. Methods and algorithms exist for community detection in social networks. In this paper, we used a Louvain community detection algorithm to effectively extract communities. In the Louvain Method of community detection, first small communities are found by optimizing modularity locally on all nodes, then each small community is grouped into one node and the first step is repeated [39].
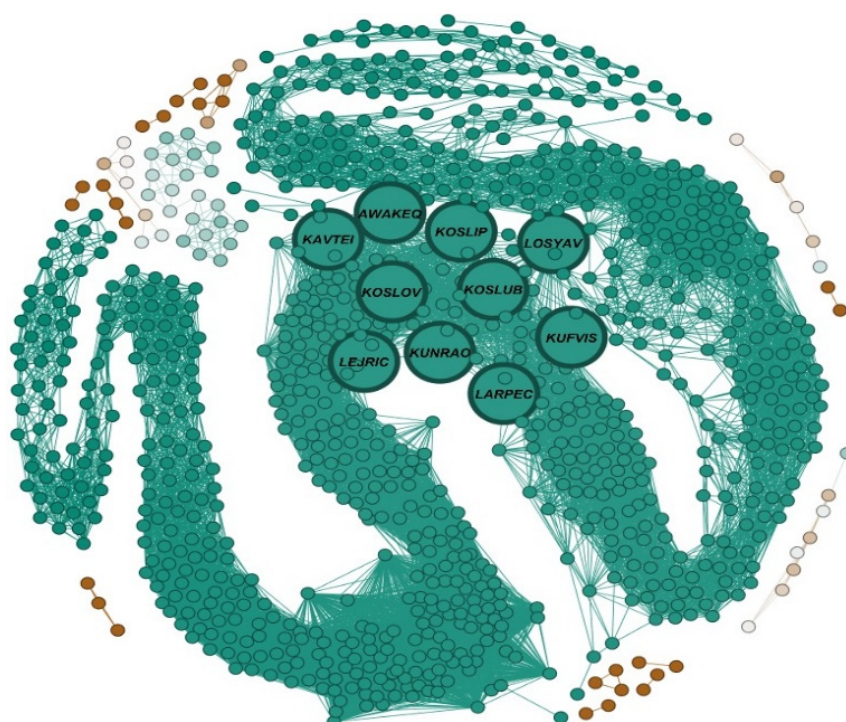

Figure 5. Visualization of the MOFSocialNet with the top 10 MOFs with the highest closeness centrality, drawn with Fruchterman Reingold layout in the Gephi software. The nodes in bold indicate the MOFs with the highest closeness centrality.

In MOFSocialNet, 24 communities were identified through the Lovain method. In Figure 6, we visualized the graphs of the 24 communities. To improve the readability of the graph, we removed the MOFs node labels. It should be noted that using the Gephi software, a further round of community detection was carried out on each of the 24 communities using the Louvain algorithm. The colors of the graph indicate the different communities. Thus, the nodes of the same color belong to the same community.

One metric to evaluate the extracted community is modularity. Modularity measures how strongly a network can be divided into different communities. Networks with high modularity have dense connections between the nodes within each community but sparse connections between nodes in different communities [39]. The modularity value is in the range of −0.5 to 1. A value of 1 indicates the highest modularity, the modularity of the MOFSocialNet with 1000 MOFs is 0.748. Table 4 lists the different MOFs communities depicted in Figure 6, indicates the number of MOF nodes in each community and the MOF with highest centrality in each community.
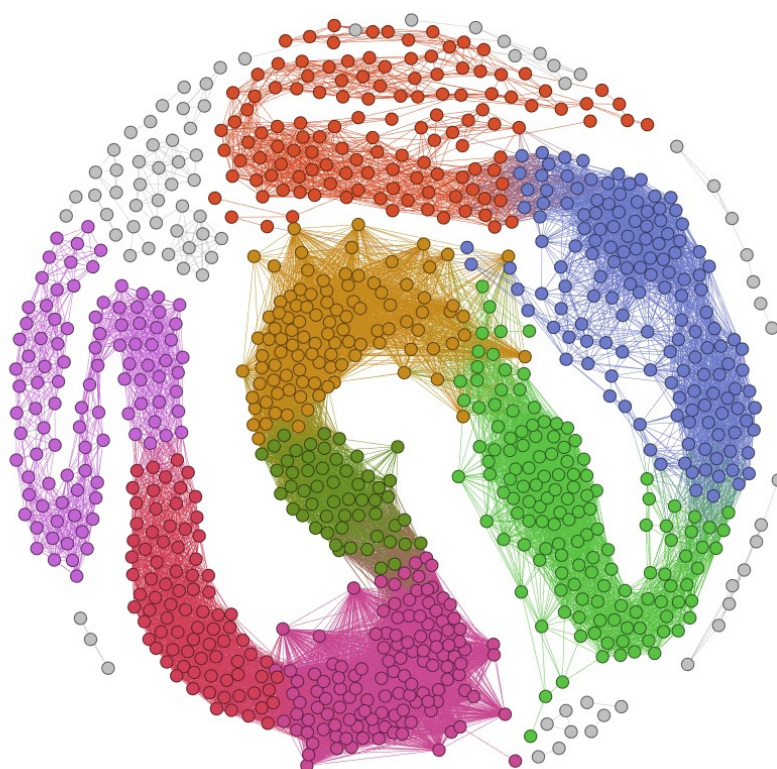
Figure 6. Visualization of MOFs communities in the MOFSocialNet. The node colors represent the community nodes, which are derived from the Lovain method. Nodes of similar color form a community. The graph is shown with the Fruchterman Reingold layout in the Gephi software.

Table 4. Community descriptions in the MOFSocialNet.

Community	MOF with Highest Centrality in Community	No. of MOF	Community	MOF with Highest Centrality in Community	No. of MOF	Community No.	MOF with Highest Centrality in Community	No. of MOF	Community	MOF with Highest Centrality in Community	No. of MOF
1	j.jlumin.2016.09.057_mmc2	54	7	KITCAT	139	13	j.matlet.2016.01.080_mmc1	2	19	KICXAX	2
2	KUYPIF	154	8	KUNRAO	103	14	JODGEP	3	20	KIYMIP	2
3	KALGEL	84	9	ACALIB	2	15	KUDWAJ	100	21	KOJZEP	2
4	ABESUX	3	10	KUVMIZ	8	16	KABXAO	7	22	KONFIE	2
5	AXIPEE	128	11	AVETAY	125	17	KARNAU	3	23	KUGCOF	2
6	LELDOX	22	12	acsami.6b11116_am6b11116_si_002	5	18	JUFREJ	10	24	JOGSAA	5

2.5. Application of MOFSocialNet to Predict the Crystal Density of Unknown MOFs

To evaluate whether properties of new MOFs can be predicted using MOFSocialNet, we randomly chose three MOFs from the CoRE MOFs database. For these MOFs, we excluded the crystal density as input during featurization and placed the MOFs within the MOFSocialNet. We then predicted the crystal density of the new MOFs by simply averaging the crystal density of the ten nearest neighbors. The results show an outstanding prediction accuracy for the crystal density of 99.69% for MOF ABAYOU, 99.79% for ABIXOZ, and 99.96% for ACOLIP. Table 5 presents the density range of the MOFs in the communities.

2.6. Use of MOFSocialNet to Predict Gas Adsorption in the Metal Organic Framework

In the final part of our investigation, we evaluated how well the communities extracted from MOFSocialNet can be exploited for predicting gas adsorption properties of MOFs for CO_2 and CH_4. To evaluate the performance, we compared the prediction performance of MOFSocialNet with three common ML models, namely K Nearest Neighbour (KNN), Gradient Boosting Regression, and Deep Learning. All ML predictions were performed using Rapidminer Machine Learning tools. The efficiency of each ML algorithm was assessed by computing with the mean absolute error (MAE) which is given in Equation (5).

$$MAE = \frac{\sum_{i=1}^{n}|y_i - x_i|}{n} \tag{5}$$

where y_i is the prediction value of gas adsorption and x_i the true value of respected gas adsorption.

The gas adsorption method by MOFSocialNet was performed by (a) creating the MOFSocialNet, as explained in the previous section, (b) extracting communities from MOFSocialNet, and (c) predicting gas adsorption in each individual community. The MAE performance parameter for each individual community is presented using the prediction of CO_2 absorption. Moreover, the overall MAE results were compared to the three main well-known algorithms. Similar to CO_2 prediction, the prediction of gas adsorption for CH_4 is presented for each community, and the overall MAE results were compared to the three main well-known algorithms. Using MOFSocialNet significantly improved the prediction compared to the reference ML algorithms, as shown in the results in Figure 7.

Table 5. Density range of the MOFs in the communities.

Community Number	Density (g/cm³) Range	
8	0.380351	0.334035
14	0.43704	0.402001
18	0.462868	0.458126
21	0.514247	0.506406
10	0.554774	0.536397
4	0.674371	0.58881
15	0.918625	0.697914
1	1.01242	0.926612
9	1.08519	1.01371
0	1.13994	1.08774
5	1.23288	1.141
20	1.35327	1.34717
11	1.42464	1.22885
12	1.42645	1.41674
6	1.68807	1.43058
7	1.78668	1.76873
13	2.17773	2.16342
3	2.25323	1.58742
22	2.25449	2.20818
16	2.5731	2.30777
23	2.84923	2.83999

Table 5. *Cont.*

Community Number	Density (g/cm^3) Range	
17	3.02612	2.73182
19	3.58491	3.50049
2	3.8283	3.69928

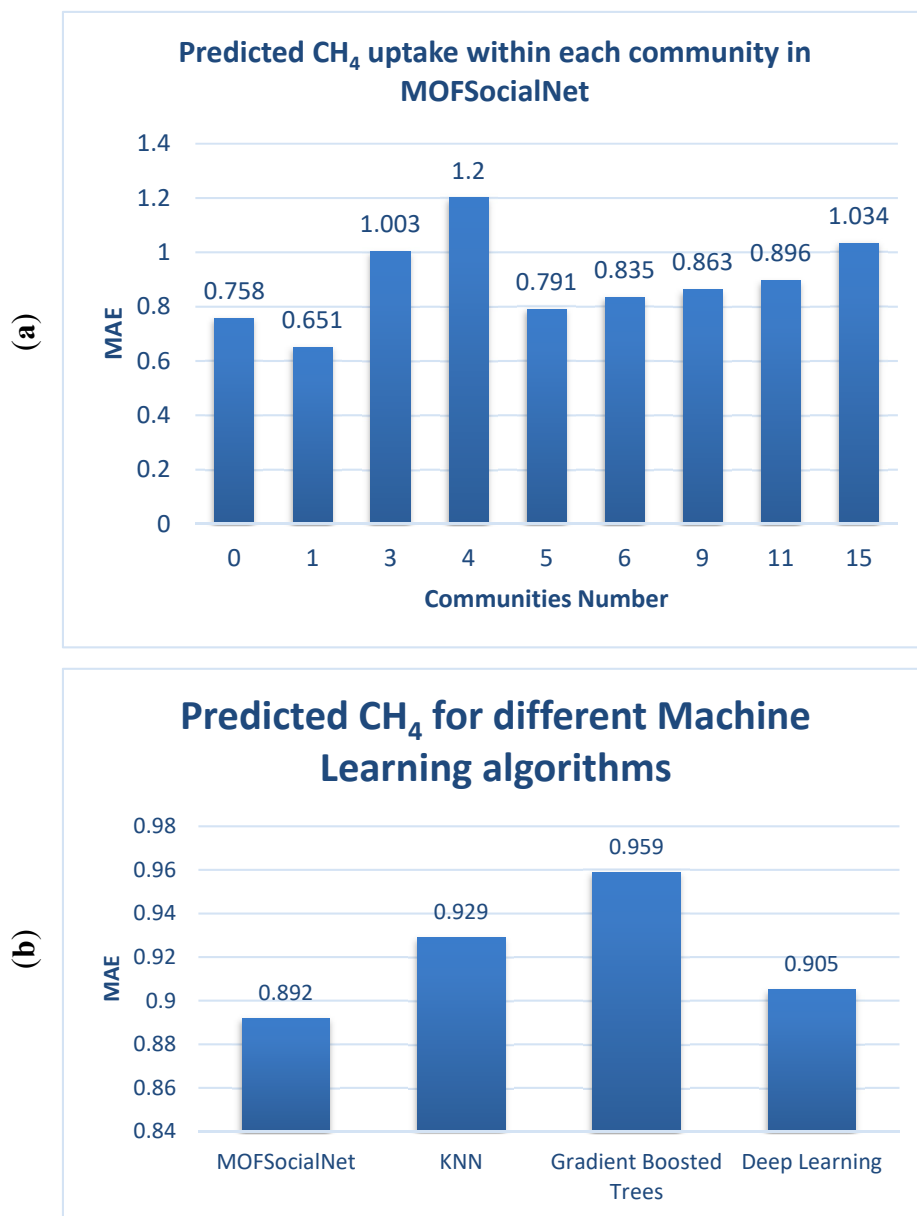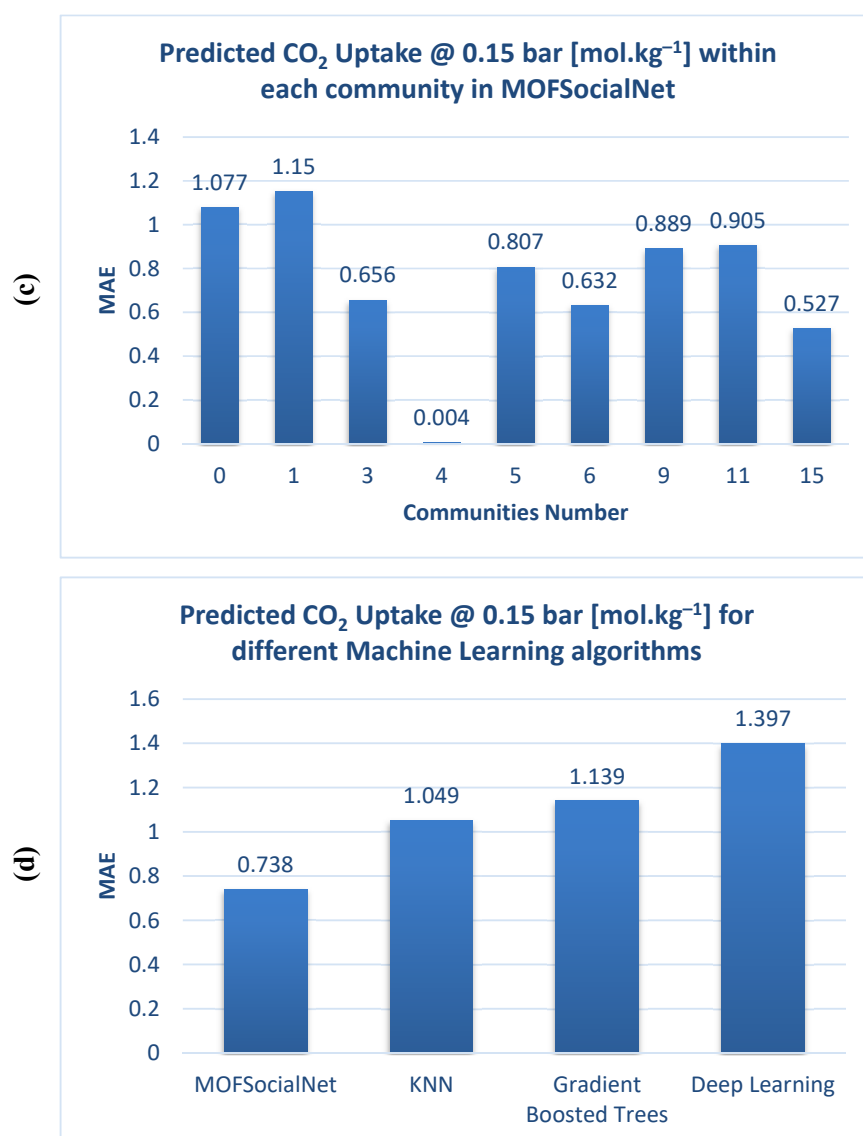The findings indicate a significant improvement in the prediction presented in Table 6.

Figure 7. *Cont.*

Figure 7. (**a**) Prediction of CH_4 adsorption within each community in MOFSocialNet; (**b**) Average prediction of CH_4 uptake for all communities in MOFSocialNet compared with the prediction of other machine learning algorithms; (**c**) Prediction of CO_2 uptake @ 0.15 bar (mol·kg^{-1}) within each community in MOFSocialNet; and (**d**) Average prediction of CO_2 uptake @ 0.15 bar (mol·kg^{-1}) for all communities in MOFSocialNet compared with the prediction of other machine learning algorithms. Only communities with a reasonable number of MOFs to predict are depicted in this figure.

Table 6. Evaluation of predictive performance of CO_2 and CH_4 using MOFSocialNet and compared reference machine learning algorithms.

Method Name	MAE for CO_2 Prediction	MAE for CH_4 Prediction
KNN	1.049	0.929
Gradient Boosted Trees	1.139	0.959
Deep Learning	1.397	0.905
MOFSocialNet	**0.738**	**0.892**

3. Conclusions and Future Direction

In this paper, we demonstrated that social network analysis (SNA), a tool developed in the social sciences, is well suited to analyse MOF structural databases. The MOFSocialNet was constructed using geometrical descriptors provided in an existing MOF database, CoRE-MOF, yielding an undirected, weighted, and heterogeneous social network. We then used MOFSocialNet as a new tool to guide MOF researchers through the vast chemical space of existing and hypothetical MOFs. For demonstration, we employed SNA to identify the most representative MOFs in this set of research data and to detect MOFs communities, i.e., families of MOFs with similar properties. Furthermore, within each community, the SNA identifies the most representative MOFs structure. Our approach can help to rationally select the most appropriate MOF structures for future studies, such as the most promising or diversified MOF. To demonstrate the feasibility of property prediction via SNA we trained communities extracted from MOFSocialNet, then predicted CO_2 and CH_4 adsorption and evaluated the accuracy of the prediction. Interestingly, SNA outperformed three common ML models, namely K Nearest Neighbour (KNN), Gradient Boosting Regression, and Deep Learning. The proposed SNA approach can accelerate the analysis of MOFs structure, even by increasing the amount of theoretical and experimental data on MOFs. MOFSocialNet as a novel framework can be extended to processing and curating large MOFs databases.

Author Contributions: Conceptualization, M.J., M.T. and C.W.; methodology, M.J., M.T. and C.W.; software, M.J., M.T. and C.W.; validation, M.J., M.T. and C.W.; formal analysis, M.J., M.T. and C.W.; investigation, M.J., M.T. and C.W.; resources, M.J., M.T. and C.W.; data curation, M.J., M.T. and C.W.; writing—original draft preparation, M.J., M.T. and C.W.; writing—review and editing, M.J., M.T. and C.W.; visualization, M.J., M.T. and C.W.; supervision, M.J., M.T. and C.W.; project administration, M.J., M.T. and C.W.; funding acquisition, M.J., M.T. and C.W. All authors have read and agreed to the published version of the manuscript.

References

1. Moghadam, P.Z.; Li, A.; Wiggin, S.B.; Tao, A.; Maloney, A.G.P.; Wood, P.A.; Ward, S.C.; Fairen-Jimenez, D. Development of a Cambridge Structural Database Subset: A Collection of Metal–Organic Frameworks for Past, Present, and Future. *Chem. Mater.* **2017**, *29*, 2618–2625. [CrossRef]
2. Haldar, R.; Mazel, A.; Krstić, M.; Zhang, Q.; Jakoby, M.; Howard, I.A.; Richards, B.; Jung, N.; Jacquemin, D.; Diring, S.; et al. A de novo strategy for predictive crystal engineering to tune excitonic coupling. *Nat. Commun.* **2019**, *10*, 2048. [CrossRef] [PubMed]
3. Jablonka, K.M.; Ongari, D.; Moosavi, S.M.; Smit, B. Big-Data Science in Porous Materials: Materials Genomics and Machine Learning. *Chem. Rev.* **2020**, *120*, 8066–8129. [CrossRef] [PubMed]
4. Chong, S.; Lee, S.; Kim, B.; Kim, J. Applications of machine learning in metal-organic frameworks. *Coord. Chem. Rev.* **2020**, *423*, 213487. [CrossRef]
5. Yan, Y.; Zhang, L.; Li, S.; Liang, H.; Qiao, Z. Adsorption behavior of metal-organic frameworks: From single simulation, high-throughput computational screening to machine learning. *Comput. Mater. Sci.* **2021**, *193*, 110383. [CrossRef]
6. Burner, J.; Schwiedrzik, L.; Krykunov, M.; Luo, J.; Boyd, P.G.; Woo, T.K. High-Performing Deep Learning Regression Models for Predicting Low-Pressure CO_2 Adsorption Properties of Metal–Organic Frameworks. *J. Phys. Chem. C* **2020**, *124*, 27996–28005. [CrossRef]
7. Batra, R.; Chen, C.; Evans, T.G.; Walton, K.S.; Ramprasad, R. Prediction of water stability of metal–organic frameworks using machine learning. *Nat. Mach. Intell.* **2020**, *2*, 704–710. [CrossRef]

8.   Mukherjee, K.; Colón, Y.J. Machine learning and descriptor selection for the computational discovery of metal-organic frameworks. *Mol. Simul.* **2021**, *47*, 1–21. [CrossRef]

9.   Luo, Y.; Bag, S.; Zaremba, O.; Cierpka, A.; Andreo, J.; Wuttke, S.; Friederich, P.; Tsotsalas, M. MOF Synthesis Prediction Enabled by Automatic Data Mining and Machine Learning. *Angew. Chem. Int. Ed.* **2022**. [CrossRef]

10. Nandy, A.; Duan, C.; Kulik, H.J. Using Machine Learning and Data Mining to Leverage Community Knowledge for the Engineering of Stable Metal–Organic Frameworks. *J. Am. Chem. Soc.* **2021**, *143*, 17535–17547. [CrossRef]

11. Pétuya, R.; Durdy, S.; Antypov, D.; Gaultois, M.W.; Berry, N.G.; Darling, G.R.; Katsoulidis, A.P.; Dyer, M.S.; Rosseinsky, M.J. Machine-Learning Prediction of Metal–Organic Framework Guest Accessibility from Linker and Metal Chemistry. *Angew. Chem. Int. Ed.* **2022**, *61*, e202114573. [CrossRef] [PubMed]

12. Li, H.; Zhang, X.; Zhao, C. Explaining social events through community evolution on temporal networks. *Appl. Math. Comput.* **2021**, *404*, 126148. [CrossRef]

13. Feng, C.; Eickhoff, S.B.; Li, T.; Wang, L.; Becker, B.; Camilleri, J.A.; Hétu, S.; Luo, Y. Common brain networks underlying human social interactions: Evidence from large-scale neuroimaging meta-analysis. *Neurosci. Biobehav. Rev.* **2021**, *126*, 289–303. [CrossRef] [PubMed]

14. García-Álvarez, M.T.; Novo-Corti, I.; Varela-Candamio, L. The effects of social networks on the assessment of virtual learning environments: A study for social sciences degrees. *Telemat. Inform.* **2018**, *35*, 1005–1017. [CrossRef]

15. Rhee, T.G.; Marottoli, R.A.; Monin, J.K. Diversity of social networks versus quality of social support: Which is more protective for health-related quality of life among older adults? *Prev. Med.* **2021**, *145*, 106440. [CrossRef]

16. Pomare, C.; Long, J.C.; Churruca, K.; Ellis, L.A.; Braithwaite, J. Social network research in health care settings: Design and data collection. *Soc. Netw.* **2019**. [CrossRef]

17. Anderson, J.G. Evaluation in health informatics: Social network analysis. *Comput. Biol. Med.* **2002**, *32*, 179–193. [CrossRef]

18. Guney, E.; Menche, J.; Vidal, M.; Barabasi, A.L. Network-based in silico drug efficacy screening. *Nat. Commun.* **2016**, *7*, 10331. [CrossRef]

19. Hohman, M.; Gregory, K.; Chibale, K.; Smith, P.J.; Ekins, S.; Bunin, B. Novel web-based tools combining chemistry informatics, biology and social networks for drug discovery. *Drug Discov. Today* **2009**, *14*, 261–270. [CrossRef]

20. Briard, L.; Ezenwa, V.O. Parasitism and host social behaviour: A meta-analysis of insights derived from social network analysis. *Anim. Behav.* **2021**, *172*, 171–182. [CrossRef]

21. Xu, B.; Guan, J.; Wang, Y.; Wang, Z. Essential Protein Detection by Random Walk on Weighted Protein-Protein Interaction Networks. *IEEE ACM Trans. Comput. Biol. Bioinform.* **2017**, *16*, 377–387. [CrossRef] [PubMed]

22. Meng, X.; Xiang, J.; Zheng, R.; Wu, F.; Li, M. DPCMNE: Detecting protein complexes from protein-protein interaction networks via multi-level network embedding. *IEEE ACM Trans. Comput. Biol. Bioinform.* **2021**. [CrossRef] [PubMed]

23. Kreft, C.S.; Angst, M.; Huber, R.; Finger, R. Social network data of Swiss farmers related to agricultural climate change mitigation. *Data Brief* **2021**, *35*, 106898. [CrossRef] [PubMed]

24. Mekonnen, D.A.; Gerber, N.; Matz, J.A. Gendered Social Networks, Agricultural Innovations, and Farm Productivity in Ethiopia. *World Dev.* **2018**, *105*, 321–335. [CrossRef]

25. Gan, C.; Voda, M.; Wang, K.; Chen, L.; Ye, J. Spatial network structure of the tourism economy in urban agglomeration: A social network analysis. *J. Hosp. Tour. Manag.* **2021**, *47*, 124–133. [CrossRef]

26. Roh, T.H. The Sharing Economy: Business Cases of Social Enterprises Using Collaborative Networks. *Procedia Comput. Sci.* **2016**, *91*, 502–511. [CrossRef]

27. Sánchez-Corcuera, R.; Bilbao-Jayo, A.; Zulaika, U.; Almeida, A. Analysing Centralities for Organisational Role Inference in Online Social Networks. *Eng. Appl. Artif. Intell.* **2021**, *99*, 104129. [CrossRef]

28. Moscato, V.; Picariello, A.; Sperlí, G. Community detection based on Game Theory. *Eng. Appl. Artif. Intell.* **2019**, *85*, 773–782. [CrossRef]

29. Wang, Y.; Cao, J.; Bu, Z.; Jiang, J.; Chen, H. Proximity-based group formation game model for community detection in social network. *Knowl.-Based Syst.* **2020**, *214*, 106670. [CrossRef]

30. Yuan, Q.; Liu, B. Community detection via an efficient nonconvex optimization approach based on modularity. *Comput. Stat. Data Anal.* **2020**, *157*, 107163. [CrossRef]

31. Yun, Z.; Liu, Y.; Li, J.; Zhu, J.; Yang, C.; Yang, W.; Wen, C. Wocda: A Whale Optimi-zation Based Community Detection Algorithm. *Phys. A Stat. Mech. Appl.* **2020**, *539*, 122937. [CrossRef]

32. Wang, S.; Gong, M.; Liu, W.; Wu, Y. Preventing epidemic spreading in networks by community detection and memetic algorithm. *Appl. Soft Comput.* **2020**, *89*, 106118. [CrossRef]

33. Golsefid, S.M.M.; Zarandi, M.H.F.; Bastani, S. Fuzzy duocentric community detection model in social networks. *Soc. Netw.* **2015**, *43*, 177–189. [CrossRef]

34. Naderipour, M.; Zarandi, M.H.F.; Bastani, S. A type-2 fuzzy community detection model in large-scale social networks considering two-layer graphs. *Eng. Appl. Artif. Intell.* **2020**, *90*, 103206. [CrossRef]

35. Opsahl, T.; Agneessens, F.; Skvoretz, J. Node centrality in weighted networks: Generalizing degree and shortest paths. *Soc. Netw.* **2010**, *32*, 245–251. [CrossRef]

36. Han, J.; Kamber, M.; Pei, J. Data Mining Concepts and Techniques Third Edition. *Morgan Kaufmann Ser. Data Manag. Syst.* **2011**, *5*, 83–124. [CrossRef]

37. Cormen, T.H.; Leiserson, C.E.; Rivest, R.L.; Stein, C. *Introduction to Algorithms*; MIT Press: Cambridge, MA, USA, 2009.

38. Newman, M. Networks: An Introduction. 2010: Oxford University Press. *Artif. Life* **2012**, *18*, 241–242. [CrossRef]

39. Blondel, V.D.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, *2008*, P10008. [CrossRef]

40. Moosavi, S.M.; Nandy, A.; Jablonka, K.M.; Ongari, D.; Janet, J.P.; Boyd, P.G.; Lee, Y.; Smit, B.; Kulik, H.J. Understanding the diversity of the metal-organic framework ecosystem. *Nat. Commun.* **2020**, *11*, 4068. [CrossRef]