Preprint of: Maulana A, Atzmueller M (2020): Centrality-Based Anomaly Detection on Multi-Layer Networks Using Many-Objective Optimization. Proc. IEEE International Conference on Control, Decision and Information Technologies, Prague, Czech Republic.

Centrality-Based Anomaly Detection on Multi-Layer Networks Using Many-Objective Optimization

Asep Maulana
Department of Cognitive Science
and Artificial Intelligence,
Tilburg University, Tilburg,
The Netherlands
Email: a.maulana@uvt.nl

Martin Atzmueller
Department of Cognitive Science
and Artificial Intelligence,
Tilburg University, Tilburg,
The Netherlands
Email: m.atzmueller@uvt.nl

***Abstract*— Anomaly detection on complex network is receiving increasing attention, e. g., for finding illegal financial transactions, or for understanding the behavior of people via analyzing social network data. This paper presents a novel method for recognizing and finding anomalies in complex networks. Specifically, it targets multi-layer social network data aiming at finding abnormal behavior of some (groups of) nodes in the network. The method starts by measuring the centrality of all nodes in each layer of the multi-layer network, continues by applying many-objective optimization with full enumeration based on minimization, and obtains the Pareto Front. Objective functions to be optimized simultaneously are the centrality of each layer in the network and thus, the number of objective function are the numbers of existing layers of a multi-layer networks. After the Pareto Front settles, the set of nodes in the Pareto Front are considered as a basis for finding the set of suspected anomaly nodes, using the novel ACE-Score. The ACE-Score is calculated by considering the centrality of a node in the $i-th$ layer, the mean of the centrality in that layer, the standard deviation, and the edge density of each layer. A high ACE-Score then indicates candidate anomalous nodes. We evaluate the approach on generated synthetic network as well as real-world complex networks, demonstrating the effectiveness of the proposed approach. A key feature of our proposed approach is its interpretability and explainability, since we can directly assess anomalous nodes with respect to the network topology.**

I. INTRODUCTION

Identifying and detecting anomalies in complex networks is an important research problem that is relevant in various contexts, e. g., for behavioral analysis in social interaction networks [2], [19], in financial transaction networks regarding fraud detection [11], or in trading networks between countries, for example, to obtain interesting insights into anomalous economical issues of specific countries [15], [16].

In this work, we aim to identify and analyze potential anomalies in complex networks using many-objective optimization. We specifically focus on a specific class of networks, i. e., multiplex/multi-layer networks. For multi-layer networks, each node is not only part of a single (layered) network but part of multiple layers of a complex network and low/high centrality of node in one layer of network might not imply low/high node centrality in another layer of the network. Optimizing more than three objective function simultaneously can be categorized as many-objective optimization, yielding the Pareto Front in high dimension.

Different layers as different objective functions can be complementary, it means that if, e. g., minimizing the centrality of one layer is also minimizing the centrality of another layer as well. In contrast, the layers can be contrary (conflicting) one to the other – meaning that the structure of centrality is very different.

For our experiments, we apply data from synthetic networks generated by Erdős-Rényi random graphs with a set of different p values in order to generate structurally different network layers. In addition, we also evaluate the proposed approach using a real-world social multi-layer network consisting of different kinds of online and offline relationships (Facebook, Leisure, Work, Co-authorship, Lunch) between the Faculty members of the Computer Science department at Aarhus University. We perform experiments on those synthetic and real-world data sets, demonstrating our proposed approach. Our contributions are summarized as follows:

1) This paper presents a novel approach for identifying a set of anomalous nodes using many-objective optimization on multi-layer networks. The optimization is based on minimizing network centrality and finding a set of less important nodes in the network.
2) Furthermore, we present a novel measure (the ACE-Score) for identifying anomaly candidates balancing criteria of connectedness, centrality and density.
3) In our experiments, we evaluate the proposed method using synthetic as well as real-world data. We generate simple to interpret network structures, and demonstrate the effectiveness of our approach in this synthetic data as well as real-world multi-layer networks.
4) One particularly appealing feature of our proposed approach is its interpretability and explainability: the results of many-objective optimization and the presented ACE-Score can directly be assessed – in a human-centered way – using simple and intuitive topological network features, also connected to the network (layers).

The rest of the paper is structured as follows: Section II discusses related work. The next Section III describes the proposed method in detail before we discuss our evaluation and results in Section IV. For the last, Section V draw the conclusion with a summary and outlines interesting directions for future work.

II. RELATED WORK

The analysis and detection of irregular (unfamiliar) or exceptional patterns, i. e., anomalies, in network-structured data is a novel research area, e. g., for identifying new and/or emerging behavior, or for identifying detrimental or malicious activities, e.g., [3], [11], [18], [27]–[29].

There are different definitions of an anomaly. According to the classical definition of [10], "an outlier is an observation that differs so much from other observations as to arouse suspicion that it was generated by a different mechanism". Adapted to anomalies in networks (represented by graphs), the general graph anomaly detection problem can be defined as follows: "Given a [. . .] graph database, find the graph objects [. . .] that are rare and that differ significantly from the majority of the reference objects in the graph" [1]. Considering networks (represented by graphs), we can focus on different types of graph objects. We can consider individual node, links/edges between nodes, or more complex substructures of nodes and/or links, respectively. Currently, in the literature there are mainly approaches for handling individual point anomalies corresponding to detecting individual nodes, c. f., [1]. However, in real-world networks the situation is typically more complex than only considering point anomalies and static graphs: for these it is difficult to capture the multi-relations of the complex heterogeneous networks. Therefore, we extend our focus from point anomalies to groups and more complex structures, i. e., towards multiplex networks.

Optimizing many-objectives functions ($\geq$ 3) simultaneously, called Many-Objective Optimization is applied for tackling many problems not only in science and engineering but in social sciences as well. In this line of research, different approaches have been developed. One of them aims to reduce the complexity, such as e. g., [15]. Other approaches are multi and many-objective optimization applied for network analysis such as [8], [16], [30], and [17]

In this paper, we adapt an approach proposed by Maulana et al. [16], [17]. They presented an approach to analyze centrality and community modularities of nodes on multi-layer networks. The first step of the approach starts by measuring the centrality of all nodes on all layers of multi-layer network. This is followed by applying many-objective optimization with full enumeration of all layers based on a *maximization problem* to find the Pareto Front. In contrast to this approach, we utilize the Pareto Front as a non-dominated solution generated by many-objective optimization for *minimization* as a basis to extract a set of anomaly candidates, i. e., a set of suspected anomalous nodes from the network.

III. METHOD

In our research and experiment, we utilized several methods from network science and optimization, which are; multi-layer networks, network centrality, edge density, many-objective optimization and networks statistical approaches. Below, we first present an overview of our approach before we introduce the specific steps in detail.

A. Algorithmic Overview

In this paper, our objective is to detect anomalous behavior of some nodes in multi-layer networks; we start by identifying a set of nodes that have some unusual behavior in terms of their interaction to other nodes. For that, we focus on the network structure, and target nodes with little connectivity or deviating network structure, i. e., nodes having a low centrality in the networks. Utilizing many-objective optimization and statistical network analysis approaches, we apply the following steps for identifying anomalous behavior of nodes in multi-layer networks.

1) First, we estimate the centrality of all nodes in each layer of the multiplex network, as well as the network density of each layer.
2) Second, we apply many-objective optimization for identifying a set of less important nodes through minimization (using node centrality); the number of layers become the number of objective function to be optimized simultaneously in order to find the Pareto Front.
3) Given the Pareto Front, we select a *candidate node* from that Pareto Front, if (1) it has no connection/link in at least one layer from the multiplex network, or (2) it has a very low centrality in a layer with very high density, or its centrality is almost zero.

B. Multi-Layer Networks

A multi-layer (or multiplex) network consists of multiple layers – modeling multiple relations. It can be defined as graphs composed of a number of n nodes and m different link sets for those nodes, which we call layers. The set of nodes is denoted by V, the sets of links are symbolized by E_l, $l \in \{1...m\}$. Furthermore, a multiplex network then can be represented formally as $G = (G_1, G_2, \ldots, G_l, \ldots, G_m)$, where $G_i = (V_i, E_i)$. A visual depiction of the network with different layers is shown in Figure 1. Next, each network G_l is represented by the adjacency matrix A_l with the elements $a^l_{ij} = W^l_{ij} > 0$, where $a^l_{ij} = W^l_{ij} > 0$, if there is a positive weight of the link between those nodes v_{il} and v_{jl}, $v_{il}, v_{jl} \in E_l$ in layer l, and $a^l_{ij} = 0$ otherwise. To simplify the formalization of weighted multiplex networks, we will consider only taking a positive integer value or zero of the link between any pair of such nodes v_{il} and v_{jl} in a layer l

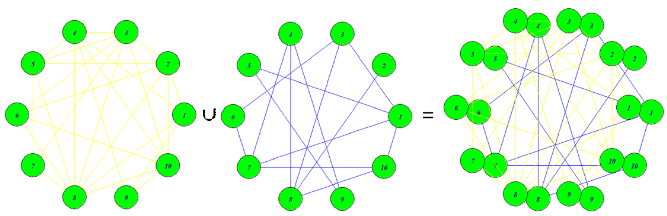

Fig. 1. Depiction sketch of a multi-layer network consist of ten nodes, with two types of different link. The difference link in this network illustrated by different color of the edge

C. Network Centrality

In network science, there are special methods for finding the most influential node/vertice in the network using the notion of so-called network centrality. Network centrality methods for finding a key player or the most influential nodes in a social setting [5], [6] include e. g.,

- Closeness centrality, which used to measure the distance from a specific node to all other nodes [22];
- Betweenness centrality, used to define the number of shortest paths past the certain nodes in the network [4];
- Eigenvector centrality, which is considers the number of links from other nodes, their importance, and to how many these nodes themselves point to, c. f., [23] and
- Degree centrality, which is center of interest on the number of peers for a connected nodes [7]; [24], respectively.

For our proposed approach, we consider eigenvector centrality, which is computed for each layer of the network. We apply eigenvector centrality, since this precisely corresponds to our intuition for estimating the notion of connections to important nodes and/or parts of the network, which is relevant for anomaly detection. However, in general, our method can be applied by using different centrality measures. Furthermore, the eigenvector centrality can be defined as follows: For a given graph $G := (V, E)$ with $|V|$ number of nodes, let $A = (a_{v,t})$, $v \in \{1, ..., |V|\}$, $t \in \{1, ..., |V|\}$ be the adjacency matrix, i.e. $a_{v,t} = 1$ if vertex v is linked to vertex t, and $a_{v,t} = 0$ otherwise. The relative centrality score of vertex v is defined as:

$$x_v = \frac{1}{\lambda} \sum_{t \in M(v)} x_t = \frac{1}{\lambda} \sum_{t=1}^{|V|} a_{t,v} x_t \,,$$

where $M(v)$ is a set of the neighbors of v while a λ is a constant. Furthermore, the eigenvector centrality in vector notation can be rewritten into a simple equation as follows: $Ax = \lambda x$; it is clearly defined that x is an eigenvector of A. Since there could be many eigenvectors of A, by convention, the eigenvector that corresponds to the largest eigenvalue is considered. There are two most important factors that influence the eigenvector centrality of each node: (1) the number of neighbors that point to the node, and (2) the weight of neighbors that point to the node. Furthermore, there is also a possibility that nodes with more neighbors have a lower eigenvector centrality compared to nodes with fewer neighbors. This condition can be happen, when the neighbors of the less connected node have higher weights.

D. Centrality-Based Many-Objective Optimization Approach

In this paper, we apply many-objective optimization of network centrality of multi-layer network and finding the Pareto Front of the less important nodes not only in a single layer but in all layers simultaneously. For each layer we thus define one objective function, since our method is applied to minimize the eigenvector centrality of that layer. By this way, for a multiplex network G with layers $G_1, ..., G_M$, then M is defined as the number of objective functions.

Each node in the each layer of the network has different centrality either it is high or low and dominated or non-dominated. The node centrality is said to be non-dominated if there is no other point which is better or equal of the centralities in different layers and better in at least one criterion or in one layer. For this approach on how to compute the non-dominated subset from a finite set of N solutions then the algorithm by Kung, Luccio and Preparata is the fastest known approach [14]. It accomplishes this task with a time complexity in $O(N \log N)$ for $M = 2, 3$ and $O\left(N (\log N)^{M-2}\right)$ for $M > 3$. For the approach proposed in this paper, we computate thus Pareto Fronts using many-objective optimization (by minimizing eigenvector centrality) to find non-dominated solutions, i. e., a set of nodes with very low importance in the multi-layer network. The set of nodes in the Pareto Front, i. e., the set of non-dominated solutions, is then the basis to select suspected anomalous nodes.

E. Statistical Analysis Based on Edge-Density

For statistical analysis of the nodes, we apply the nodes' mean centrality and standard deviation, for a respective layer, for deriving a specific score as discussed below. Furthermore, we include another parameter, i. e., the edge density: It measures the fraction of present to possible edges: In a dense graph the number of edges is close to the maximal number of edges. In contrast, a graph with only a few edges is then a sparse graph. For undirected simple graphs, the graph density is defined as: $D = \frac{2|E|}{|V|\,(|V|-1)}$. For a directed graph, this graph density is divided by two, in order to take the directionality of edges into account.

F. Assessing Anomalous Nodes

For the final step in finding a set of nodes that have no connection or have very low centrality in a high density layer, it is necessary to know the density of the layer and less important nodes which are far from the mean in that layer. We propose a special score to be applied for indicating a node as an anomaly, the so-called ACE-Score (*A*nomaly candidate based on *C*entrality *E*valuation). The ACE score considers the mean and standard deviation of the centrality of all nodes in a specific layer (of the multiplex network), e. g., one which has very high edge density compared to other layers. Using the ACE-Score for the set of nodes on the Pareto Front, we choose those with the highest ACE-Score, or those that are very close to this score and are far from the mean μ in each layer. Then, these nodes are indicated as anomalous *candidate nodes*. Also, these candidates can then further be assessed and inspected in a human-centered approach.
The ACE-Score for node v in layer l is formalized as:

$$\mathrm{ACE}(v, l) = \left(\frac{\mu_l - x_{vl}}{\sigma}\right) D_l$$

in which;
μ_l = mean of nodes centrality in layer l
x_{vl}= centrality of each node/vertex v in layer l
σ = standard deviation of nodes' centrality in layer l
D_l = edge density of network in layer l

IV. CASE STUDY AND IMPLEMENTATION

For our evaluation, we provide experiments and case studies using a synthetic as well as a real-world network data. For the synthetically generated network, we apply Erdős and Rényi models since these are standard models that are simple to interpret. For the real-world network, we utilize a standard multiplex network dataset from Aarhus university.

A. Case Study I: Synthetic Multi-Layer Networks

The first step to interpret results of the many-objective centrality optimization and statistical network analysis for anomaly detection, we started out by computing the exact Pareto Fronts of eigenvector centrality using synthetically generated artificial multiplex networks.

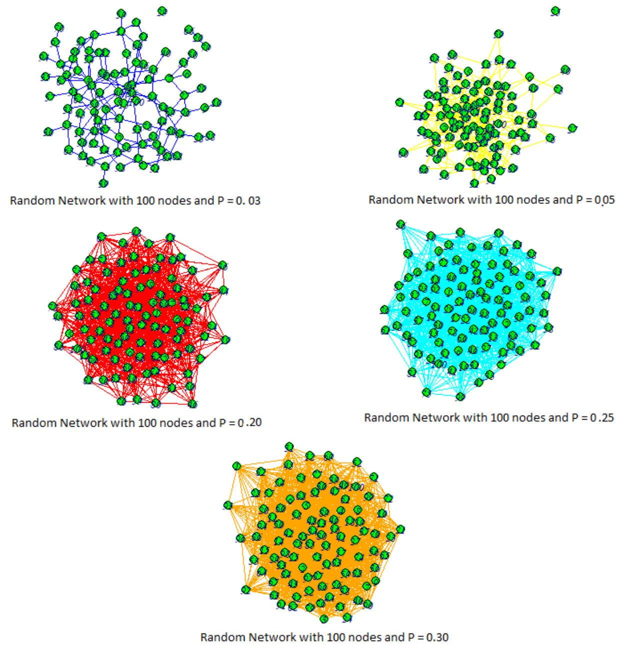

Fig. 2. Synthetic Network generated by a Erdős Rényi random graph generation strategy. Each layer consists of 100 nodes. For the different layers (1-5), layer 1 top left, layer 2 top right, etc. to bottom layer 5 we apply different link probabilities p for generation. That means, that each layer has a different edge-density.

The networks are generated via a random graph according to the Erdős and Rényi model. In this complete graph model, each edge has a probability of $1 - p$ need to be removed from the network. Therefore, we can vary the probability and number of nodes to generate different networks and layers, specifically. We generated the network with number of nodes ($N = 100$) and link probability ($p = 0.05$ to $p = 0.3$) and the number of layers consist of 5 layers. We denote the layer with g_1 to g_5. We choose five layers of this multiplex network, and apply many-objective optimization with these layers in order to compare the result of the approach with real world social multiplex network which has five layers.

The applied synthetic network based on Erdős-Rényi graph generation can be seen in Figure 2. Overall, we performed a round of experiments, and take this specific instantiation as an example for demonstrating the methodological approach. Other instantiations showed similar results.

For our example case, as a result of the estimation of eigenvector centrality and the application of many-objective optimization for synthetics networks, we found 38 nodes in the Pareto Front. From these 38 nodes, we found some nodes with very high ACE-Score, or nodes that are very close to these and show a large deviation from the mean μ. For finding nodes with very high ACE-Score we consider to extract them from the layers that have high density D. In our case, these are the layers 3, 4 and 5. Here, we found 6 nodes (35, 40, 48, 66, 68 and node 99) which are then categorized as anomalous nodes. As can be seen in Figure 3, we select the nodes with the highest ACE-Score that are also in the Pareto Front, focussing on the denser layers, while the low-density layers are neglected since these also feature quite low ACE-Scores. Intuitively, this makes sense, since a node with a low connectivity and/or high Eigenvector centrality on a layer with higher density then gets a higher chance for proposed as an anomalous node. It is important to note, that layers 1 and 2 feature some unconnected nodes - which can then be indicated as anomaly candidates for these layers based on a connectivity criterion. Such a (simple) criterion can of course be implemented complementing the Pareto-front-based selection.

TABLE I

STATISTICAL CHARACTERIZATION: SYNTHETIC MULTIPLEX NETWORK

Layer	μ	σ	D
L1 ($p = 0.03$)	0.194114	0.1911789	0.028889
L2 ($p = 0.05$)	0.34386	0.1979547	0.056566
L3 ($p = 0.20$)	0.674128	0.1472402	0.198384
L4 ($p = 0.25$)	0.651931	0.1232115	0.235556
L5 ($p = 0.30$)	0.726498	0.1220831	0.29899

	L1		L2		L3		L4		L5	
No	ACE_Score	Node	ACE_Score	Node	ACE_Score	Node	ACE_Score	Node	ACE_Score	Node
1	0.003489	7	0.029332	22	0.39361	68	0.520676	40	0.576491	35
2	0.003489	12	0.029332	70	0.37581	66	0.398917	39	0.556781	48
3	0.003489	15	0.028154	57	0.367554	99	0.38098	60	0.512226	49
4	0.003489	19	0.026904	41	0.36481	77	0.373612	85	0.49261	71
5	0.003489	20	0.025192	49	0.363811	71	0.357277	27	0.364353	75
6	0.003489	21	0.024367	96	0.362636	78	0.339156	3	0.332967	14
7	0.003489	22	0.024012	21	0.354883	33	0.290516	93	0.327641	64
8	0.003489	26	0.023848	50	0.300521	8	0.278613	78	0.323315	43
9	0.003489	33	0.023426	87	0.298178	39	0.270124	47	0.303674	30
10	0.003489	35	0.023066	78	0.297861	32	0.26853	18	0.301767	98

Fig. 3. ACE-Score for Synthetics network consist of 5 layers. In each layer, the nodes with the highest ACE-Score (in dense layers) lead to being indicated as anomalous nodes.

B. Case Study II: Real-World Multi-Layer Network

The applied network data consists of five layers modeling online and offline relationships (Facebook, Leisure, Work, Co-authorship, Lunch) between the employees of the Computer Science department at Aarhus University. The network includes 61 nodes (members at the department).

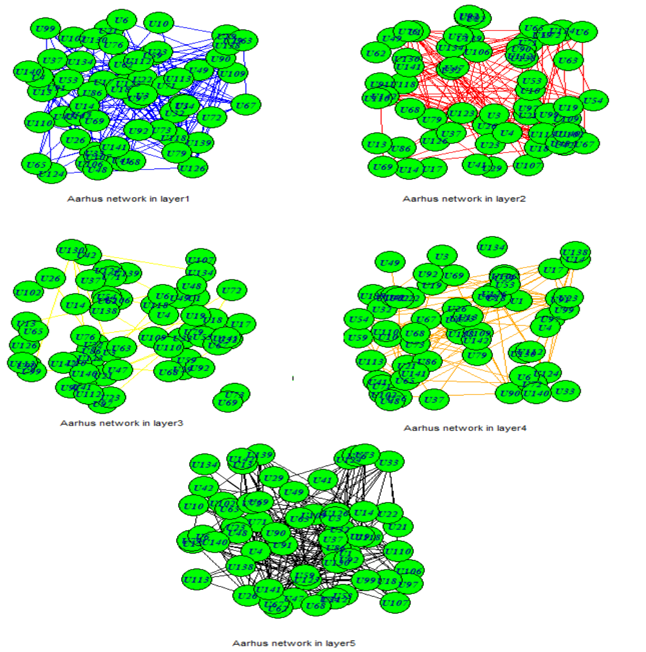

Fig. 4. Aarhus multi-layer network in which the link between nodes are different for each layer and the difference link of the layer illustrated by different color of connected edge .

As result of estimating the centrality for all nodes in all layers and applying many-objective optimization through minimization, we found the Pareto Front that consists of 6 nodes. As shown in Figure 6 these are the nodes 1, 8, 12, 14, 37 and 60. In determining the nodes to be categorized as an anomaly node, we consider high density D layers as a basis to find nodes with the highest ACE-Score. Also, we consider the ones closest to the highest ACE-Score and far from the Mean μ, being contained in the Pareto Front.

TABLE II
STATISTICAL CHARACTERIZATION: AARHUS MULTIPLEX NETWORK

Layer	μ	σ	D
Layer 1	0.257986	0.291311	0.1054645
Layer 2	0.273081	0.331148	0.06775956
Layer 3	0.073182	0.240734	0.01147541
Layer 4	0.176249	0.22261	0.04808743
Layer 5	0.266831	0.20674	0.1060109

From those nodes in the Pareto Front, we found the highest ACE-Score in each layer (this can be seen from the Figure 7), which is in the layer 1, the highest score achieved by node 60, and in layer 2 are node 1 and node 14, in layer 3 are node 4, 8, 12, 14 in layer 4 are node 1 and node 14 and in layer 5 is node 1. Therefore, in this network we can actually collect "evidence" towards designating nodes as anomalous from several layers of the network. As a consequence, we can identify four nodes out of six nodes in the Pareto Front to be categorized as anomalous nodes. Therefore, the designated final anomalous nodes are 1, 12, 14 and node 60. In the social context this then indicates, for example, that node 1, node

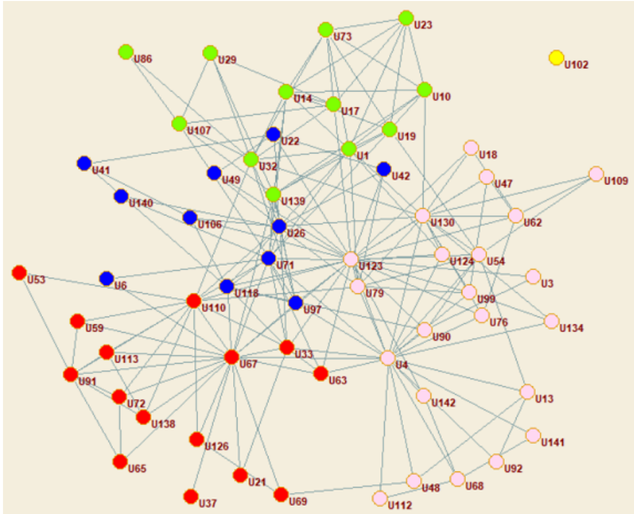

Fig. 5. Visualization of Aarhus network on layer 5. Here can be seen how the nodes connect to the others with many connection, few connection or even without any single connection.

Aarhus Pareto Front								
No	L1	L2	L3	L4	L5	Node	Label	.level
1	0.0206	0	8.12E-17	0	1.56E-17	1	U102	1
2	0.00922	0.32766	0	0.123	0.48735	8	U1	1
3	0.00873	0.25104	0	0.123	0.13253	12	U29	1
4	0.01993	0	0	0	0.08774	14	U41	1
5	0.00272	0.3851	0	0.191	0.25638	37	U10	1
6	3.06E-17	0	8.12E-17	0	0.14151	60	U140	1
7	0.04938	0.28375	0	0	0.22058	4	U106	2
8	0.02199	0	0	0	0.459	6	U118	2
...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...
56	0.41465	0.9133	8.12E-17	0.017	0.70072	44	U4	5
57	0.33325	0.7789	8.12E-17	0.1	0.65743	51	U67	5
58	0.98973	1	8.12E-17	0.456	0.48118	34	U79	6
59	0.23158	0.85764	1	1	0.18967	23	U91	6
60	0.87278	0.76229	8.12E-17	0.665	0.34929	31	U54	6
61	0.23263	0.65021	1	0.589	0.53389	46	U110	6

Fig. 6. Pareto Front of Aarhus multi-layer network, consisting of 6 nodes of non-dominated solutions. These are node 1, 8, 12, 14, 37, and node 60.

12, node 14 and node 60 are not very active in the interaction with their colleague in terms of special interactions (Layer1 = Facebook, Layer 2 = Leisure, Layer 3 = Work, Layer 4 = Co-authorship, Layer 5 = Lunch): node 60 and node 12 are very rare in interaction with their colleagues on Facebook, node 1 and node 14 are not very active in communication with their colleague related to leisure and co-authorship. The analysis process described above provides for a human-centered interpretable and explainable approach since simple network metrics and topological indicators can be easily inspected in context on the multiplex network, making use of the visualization of different layers and contextual inspection of the parameters on those [21], [26].

No	Aarhus_L1		Aarhus_L2		Aarhus_L3		Aarhus_L4		Aarhus_L5	
	ACE_Score	Node	ACE_Score	Node	ACE_Score	Node	ACE_Score	Node	ACE_Score	Node
1	0.0934	60	0.05588	1	0.00349	4	0.038073	1	0.136824	1
2	0.09254	2	0.05588	2	0.00349	6	0.038073	2	0.106959	22
3	0.09241	37	0.05588	3	0.00349	8	0.038073	3	0.103839	43
4	0.0919	41	0.05588	6	0.00349	10	0.038073	4	0.100521	49
5	0.09165	42	0.05588	10	0.00349	11	0.038073	6	0.096591	53
6	0.09097	5	0.05588	11	0.00349	12	0.038073	7	0.096591	61
7	0.09024	12	0.05588	14	0.00349	13	0.038073	14	0.095684	59
8	0.09015	45	0.05588	18	0.00349	14	0.038073	18	0.093128	56
9	0.09007	20	0.05588	20	0.00349	26	0.038073	21	0.091833	14
10	0.09006	8	0.05588	22	0.00349	27	0.038073	22	0.091408	54

Fig. 7. ACE-Score for Aarhus network (5 layers). In each layer, the nodes with the highest ACE-Score (in dense layers) lead to being indicated as candidates for anomalous nodes.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed an approach for centrality-based anomaly detection on multi-layer networks using many-objective optimization. We presented a novel method for identifying anomaly candidates minimizing Eigenvector centrality in order to obtain a Pareto Front of potential anomaly candidates. Given these, we apply a statistical criterion – formalized in the novel ACE-Score – which balances, connectivity and Eigenvector centrality of a node in relation to density relative to the containing layer.

For our evaluation, we conducted experiments using synthetic as well as real-world multi-layer network data. Specifically, we utilized an artificial dataset generated by an Erdős Rényi random graph generating approach. Furthermore, we applied a multiplex network capturing different (social) relations at the Computer Science department of Aarhus university. Overall, our results indicate, that we can quite well identify anomalous nodes based on the criteria of connectivity and importance (as estimated by the Eigenvector centrality), always relative to the density of the containing network/layer. Also, a specific advantage of the proposed method is its interpretability and explainability, referring to the network structure and topological features which can be intuitively assessed in a human-centered approach, c. f., [21].

For future work, we aim to extend the analysis towards further real-world complex networks, in order to capture and investigate further real-world phenomena about potential anomalies, e. g., in feature-rich networks [12]. In addition, we plan to analyze other centrality measures in order to compare those results with Eigenvector centrality in terms of detection performance. Furthermore, we aim to investigate interactive explanation methods for anomaly detection, also including declarative approaches, e. g., [9].

REFERENCES

[1] Akoglu, L.,Tong, H., and Koutra, D. (2015) Graph Based Anomaly Detection and Description. DMKD, 29(3), 626–688.

[2] Atzmueller, M. (2014) Data Mining on Social Interaction Networks. Journal of Data Mining and Digital Humanities 1.

[3] Atzmueller, M. (2019) Onto Model-based Anomalous Link Pattern Mining on Feature-Rich Social Interaction Networks. Proc. WWW 2019 (Companion). IW3C2 / ACM. New York, NY, USA.

[4] Barthelemy, M. (2004) Betweenness Centrality in Large Complex Networks. The European Physical Journal B, 38(2):163-168.

[5] Bonacich, P. Lloyd, P. (2001) Eigenvector-Like Measures of Centrality for Asymmetric Relations. Social Networks, 23(3), 191-201.

[6] Borgatti, S. P. and M. G. Everett (2006) A Graph-Theoretic Perspective on Centrality. Social Networks, 28(4):466-484, 2006.

[7] Freeman, L. C. (1979) Centrality in Social Networks Conceptual Clarification. Social Networks, 1(3):215-239.

[8] Gunasekara, R. C., Mehrotra, K., Mohan, C. K. (2015) Multi-Objective Optimization to Identify Key Players in Large Social Networks. Social Network Analysis and Mining, 5(1), 1-20.

[9] Guven, C., and Atzmueller, M. (2019) Applying Answer Set Programming for Knowledge-Based Link Prediction on Social Interaction Networks. Frontiers in Big Data.

[10] Hawkins, D. (1980) Identification of Outliers, Chapman & Hall, UK.

[11] Huang, D., Mu, D., Yang, L., and Cai, X. (2018) CoDetect: Financial Fraud Detection with Anomaly Feature Detection. IEEE Access, 6, 19161-19174.

[12] Interdonato, R., Atzmueller, M., Gaito, S., Kanawati, R., Largeron, C., and Sala, A. (2019) Feature-Rich Networks: Going Beyond Complex Network Topologies. Applied Network Science 4(4)

[13] Knowles, J., Corne, D. (2007) Quantifying the Effects of Objective Space Dimension in Evolutionary Multiobjective Optimization. In Evolutionary Multi-Criterion Optimization, 757-771. Springer.

[14] Kung, H. T., Luccio, F., and Preparata, F. P. (1975) On Finding the Maxima of a Set of Vectors. Journal of the ACM, 22(4), 469-476.

[15] Maulana, A., Jiang, Z., Liu, J., Bäck, T., Emmerich, M. (2015). Reducing Complexity in Many-Objective Optimization using Community Detection. In Proc. Evolutionary Computation, 3140-3147. IEEE.

[16] Maulana, A., Gametto,V., Garlaschelli.D, Yevesyeva,I., Emmerich,M. (2016) Modularities Maximization in Multiplex Network Analysis Using Many-Objective Optimization. IEEE Symposium on Computational Intelligence (SSCI), Athen.

[17] Maulana, A., and Emmerich, M. T. (2017) Towards Many-Objective Optimization of Eigenvector Centrality in Multiplex Networks. Proc. IEEE International Conference on Control, Decision and Information Technologies (CoDIT), 0729-0734. IEEE.

[18] Mittal, R., and Bhatia, M. P. S. (2018) Anomaly Detection in Multiplex Networks. Procedia Computer Science, 125, 609-616.

[19] Mitzlaff, F., Atzmueller, M., Stumme, G., and Hotho A. (2013) Semantics of User Interaction in Social Media. Complex Networks IV, Studies in Computational Intelligence 476, Springer.

[20] Marsden, P. V. (2002) Egocentric and Sociocentric Measures of Network Centrality. Social Networks, 24(4):407-422.

[21] Masiala, S., and Atzmueller, M. (2018) First Perspectives on Explanation in Complex Network Analysis. Proc. BNAIC. Jheronimus Academy of Data Science, Den Bosch, The Netherlands.

[22] Okamoto, K., Chen, W. and X.-Y. Li (2008) Ranking of Closeness Centrality for Large-Scale Social Networks. Frontiers in Algorithmics, 186–195, Springer.

[23] Page, L., Brin, S., Motwani, R. and T. Winograd (1999) The Pagerank Citation Ranking: Bringing Order to the Web.

[24] Ruhnau, B. (2000) Eigenvector-Centrality a Node-Centrality? Social Networks, 22(4), 357-365.

[25] Saxena, D., Duro, J. , Tiwari, A., Deb, K., Zhang, Q. (2013) Objective Reduction in Many-Objective Optimization: Linear and Nonlinear Algorithms. IEEE Trans. on Evolutionary Computation, 17(1),77-99.

[26] Shneiderman, B. (1996) The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. Proc. IEEE Symposium on Visual Languages, 336–343. IEEE.

[27] Wang, J. F., Liu, X., Zhao, H., and Chen, X. C. (2018) Anomaly Detection of Complex Networks based on Intuitionistic Fuzzy Set Ensemble. Chinese Physics Letters, 35(5), 058901.

[28] Yasami, Y. (2018) Anomaly Detection in Dynamic Complex Networks. Modern and Interdisciplinary Problems in Network Science: A Translational Research Perspective, 239.

[29] Zhao, Y., Chen, J., Wu, D., Teng, J., and Yu, S. (2019) Multi-Task Network Anomaly Detection using Federated Learning. Proc. International Symposium on Information and Communication Technology, 273-279.

[30] Zeng, Y., and J. Liu. (2015) Community Detection from Signed Social Networks using a Multi-Objective Evolutionary Algorithm. Proc. Asia Pacific Symposium on Intelligent and Evolutionary Systems. Springer.

ACKNOWLEDGMENT

We gratefully acknowledge support by the German Research Foundation (DFG, project MODUS) under grant AT 88/4-1.