Multi-Layer Networks: Origin, Community Detection, Applications

Babak Farzad, Oksana Pichugina, Liudmyla Koliechkina

Abstract— Communities are common structures in social networks. These structures, typically, are formed by different attributes and consequently have different textures in the network. Standard Community Detection (CD) methods detect and extract them to some degree: they often form a node partition clearly related to a dominant node attribute. Such methods are unable to detect the whole variety of communities in the network. We study CD on multi-attributed affiliation networks that are networks with nodes decorated by a number of attributes and edges forming due to their similarity. The networks are represented as a composition of single node attribute networks called one-layer networks and yield node partitions into the attribute clusters. We believe that these partitions can be detected by standard CD algorithms applied to a network accumulated both structural information and node attributes. We propose an iterative method called Multi-Layer Community Detection Algorithm (MLCDA) including two stages a synthesis phase of utilizing available network data and a decomposition phase in which communities are extracted layer by layer. The synthesis includes the conversion of an original network into a weighted one based on assumptions about the network model, construction of an association network accumulated the node attributes, and synthesis of the networks in an accumulated network. In the decomposition phase, CD is conducted on the accumulated network; for the obtained partition an underlying node attribute is determined; an approximation network for the corresponding one layer network is constructed and extracted from the accumulated network; these steps are repeatedly repeated.

Keywords— Social Networks, Community Detection, Attributed Networks, Node Partition, Accumulated Networks

I. Introduction

Network Analysis is an area of research that has been studied intensively lately [3-5, 10, 12, 31]. Among a variety of networks, social networks have been of particular interest [9,14-16,27,29,36].

Babak Farzad is Associate Professor of Mathematics, Mathematics and Statistics Department, Brock University, 1812 Sir Isaac Brock Way, St. Catharines, ON L2S 3A1, Canada, (e-mail: bfarzad@brocku.ca)

Oksana Pichugina is Associate Professor of Applied Mathematics, Computer Science Department, National Aerospace University "Kharkiv Aviation Institute", 17 Chkalova str., Kharkiv, 61070, Ukraine (e-mail: oksanapichugina1@gmail.com)

Liudmyla Koliechkina is Professor of Mathematics, Department of Informatics in Economic Systems, Poltava University of Economics and Trade, 3 Koval St. Poltava, 36014, Ukraine, and collaborates with Algorithms and Databases Department, University of Lodz, 3 Uniwersytecka Street, Lodz, 90-137, Poland (e-mail: ludaplt1971@gmail.com)

Researchers investigate the structural characteristics of different networks, network formation models, and many other relevant questions.

In many types of networks (e.g., social networks) observable and tightly bound groups of elements called *communities* exist [8,16,36]. How and why communities arise are interesting and important questions.

It is also important to find efficient algorithms for detecting such communities. Researchers in *Community Detection* (CD) develop *CD Algorithms* (CDAs) with the purpose of extracting communities quickly and qualitatively for networks that nowadays can be immeasurably large and complex.

II. Background and Motivation

Assume that our goal is to study the community structure of a social network (of people) and that we have complete information about every individual including their connections and the strength of each connection. Then the output of a typical CDA will be predictable: it would be a division by families because family ties are (on average) very strong. Thus, we fail to capture communities formed by friends or colleagues.

Our approach is to detect and eliminate the dominant subnetworks (e.g., the family relationships) that create the first layer of the global network. Conducting CD in the remaining network would demonstrate a new community structure. The result, however, is not predictable - for some people, the next important thing is friendship, for others their hobby or their job and so on. Nevertheless, it makes sense to detect the communities formed by another factor and eliminate it from the consideration, so that other new layers of the global network can be detected. These factors of the nodes are called their *attributes*, and such networks are called *attributed networks.* The network built based on the similarity of the node attributes is called an *association network.* Due to the variety of characteristics, nodes (and edges) in social networks are heterogeneous. For the described multi-layer community detection, the heterogeneity and in particular the presence of node attributes, are crucial.

In this article, we study community detection in attributed networks that combines a network structure accumulated in edges data and node attribute information.

III. Definitions and Notations

Attributed Networks Any social network can be represented as follows:

Definition 1 *[13] A social network is a hybrid graph, which is represented in the form:*

$$G=(V,E,\Lambda,\Lambda'),\qquad(1)$$

where V is the set of nodes (the social network's users), E *is the set of edges (these users relationships), A and A' contain information about attributes related to each node $v\in V$ and each edge $\{u,v\}\in E$, respectively.*

The network represented in the form (1), where $\Lambda\cup\Lambda'\neq 0$, is an *attributed network.* So, an attributed network representing individuals' relationships is social. Let $n=|V|$, $m=|E|$, and *K*, *K'* denote the number of node's and edge's attributes, respectively. An attributed network is a *multi-node attributed network* (MNA-network) if *K* > 1.

Our focus will be on MNA-networks.

Let us introduce some notations for MNA- networks whose attributes take a finite number of values:

$J_K=\{1,...,K\}$, $\overline{AT}^n=(AT^{nk})_{k\in J_K}$ is a tuple of discrete nodes' attributes (ATNs) taking values $at^{nk}=\{at_l^{nk}\}_{l\in J_{L_k}},k\in J_K$ (ATNVs); $\overline{AT}^e=(AT^{ek})_{k\in J_{K'}}$ is a tuple of discrete edges' attributes (ATEs) taking values

$$at^{ek}=\{at_l^{ek}\}_{l\in J_{L'_k}},k\in J_{K'}\ \text{(ATEVs)}.$$

$\Lambda=(a_i^k)_{v_i\in V,k\in J_K}$, $a_i^k\in at^{nk}$ - a value of AT^{nk} of the node v_i; $\Lambda'=(a'^k_{\{i,j\}})_{\{v_i,v_j\}\in E,k\in J_{K'}}$, $a_{\{i,j\}'k}$ - a value of AT^{ek} of the edge $\{v_i,v_j\}$. Let

$$\begin{aligned}&\bar{a}_i=(a_i^k)_{k\in J_K},\bar{a}'_{\{i,j\}}=(a'^k_{\{i,j\}})_{k\in J_{K'}}\\&(a_i^k\in at^{nk},a'^k_{\{i,j\}}\in at^{ek})_k\end{aligned}\qquad(2)$$

Denote the tuples of ATNVs of a node v_i and ATEVs of an edge $\{v_i,v_j\}\in E$, respectively ($v_i,v_j\in V$). In terms of (2), *A* and *A'* in (1) can be represented as follows:

$$\Lambda=(\bar{a}_i)_{i\in J_n},\Lambda'=(\bar{a}'_{\{i,j\}})_{\{i,j\}\in E}.$$

Assume that E is formed based on similarity of node attributes (e.g., in an affiliation network), then

$$\text{if }\forall k\, a_i^k\neq a_j^k\Rightarrow\{v_i,v_j\}\notin E.\qquad(3)$$

So *G* is composed of auxiliary networks

$$\bar{G}=\{G^k=(V,E(G^k))\}_{k\in J_K}$$

related to single node attributes:

$$G=\cup_k G^k,\qquad(4)$$

and for each $E(G^k)$ the following holds:

$$\text{if } a_i^k\neq a_j^k\Rightarrow\{v_i,v_j\}\notin E(G^k).$$

Each G^k is considered as a layer of G formed by the attribute AT^k, and we refer to it as a *one-layer network.*

If in an attributed network, edges exist only between vertices with sufficient level of their node attributes association, then such network is called an association network [10]. For the case of discrete node attributed network G, the sufficient level of association between any pair of nodes having identical attributes is established. The *association network* G^a corresponding to G is a network with edges between any two nodes sharing a common node attribute.

Therefore, in contrast with (3), for an association network, the edge condition is the following:

$$\text{if }\exists k\, a_i^k=a_j^k\Rightarrow\{v_i,v_j\}\in E.\qquad(5)$$

From (5) can be seen as a decomposition of type (4)

$$G^a=\cup_k G^{ak},\qquad(6)$$

where $\bar{G}^a=\{G^{ak}=(V,E(G^{ak}))\}_k$ are G^a subnetworks related to single node attributes.

Networks Operations The *adjacency matrix* of network G is a matrix A of order n such that

$a_{ij}=1$ if v_i,v_j are adjacent, and 0, otherwise.

The network's *weighted adjacency matrix* (WAM) is a matrix A^w of weights of its edges. A *linear combination of the networks* $\{G_i=(V,E(G_i))\}_i$ is a network with a WAM given by the corresponding linear combination of weighted adjacency matrices (WAMs) $\{A_i\}_i$ of the networks.

Let $C\subseteq V$, *G[C]* be an *induced G-subnetwork* by $C: G[C]=(C,E(C))$, $n(C)=|C|$. Let $||b||=\sum_{i\in J_n}b_i$ denote the norm *of a vector* $b\in R^n$. If $\alpha\in R^K:\alpha\geq 0,||\alpha||=1$ then the linear combination $\sum_{i\in J_K}\alpha_i\cdot G_i$ of $\{G_i\}_{i\in J_K}$ is called the *weighted network sum.*

The sum $\omega(G)$ of *E*-weights is a *weight* of the network *G:* $\omega(G)=\sum_{i,j\in J_n}a_{ij}^w$. A network *G* of the weight one is called a *normalized network*:

$$\omega(G)=1.\qquad(7)$$

A *network partition* is a partition of its nodes $\mathcal{C}=\{C_l\}_{l\in J_L}$. We will consider two types of the partitions - node partitions into communities (obtained as a result of implementing a CDA) and node partitions into clusters related to different values of a particular node attribute. Then node covers are formed from these partitions (see Section 4.b). For the clarification we use the following notation for a partition $\mathcal{C}^*=\{C_{l^*}\}_{l^*\in J_{L^*}}$ into communities:

$$|C_{l^*}| = n_{l^*}, l^* \in J_{L^*}, \sum_{l^* \in J_{L^*}} n_{l^*} = n. \qquad (8)$$

IV. Related Work

A. Association Network Inference Problem

Normally in a group of people, new connections are formed more frequently with people that have common hobbies and interests. At the same time, people with common interests are not necessarily friends. However, in association networks, an edge exists if and only if the level of association is attained. If we know a function of the similarity $sim(i,j) = F(\bar{a}_i, \bar{a}_j)$, where $\bar{a}_i, \bar{a}_j$ - are vectors of attributes of nodes $v_i, v_j \in V$ (see (2)), then given a level α of association, an edge set $E(G^a)$ of an association network G^a can be formed according to the rule: $\{v_i, v_j\} \in E(G^a)$ if $F(\bar{a}_i, \bar{a}_j) \geq \alpha$.

Unfortunately, function F(.) is usually unknown. So typically the following assumptions are made when available network information is sufficient to restore missing elements of the network. For example, [10] focuses on a case where the information, contained in node attributes, is sufficient to establish links. In particular, as a similarity function, one may take the correlation between vectors of node attributes

$$\bar{a}_i, \bar{a}_j \text{ - } sim(i,j) = \rho_{\bar{a}_i, \bar{a}_j}$$

and establishing edges are suggested based on the results of verifying the following statistical hypotheses - $H_0^{ij}: \rho_{\bar{a}_i, \bar{a}_j} = 0$ versus $H_1^{ij}: \rho_{\bar{a}_i, \bar{a}_j} \neq 0$.

Since links in association networks are formed based only on a similarity of node attributes, the algorithm is also an approach to solve Networks Inference of Association Networks Problem. A drawback of the technique [10] is that the correlation does not consider the weights of node attributes. Moreover, this approach does not work when only one node attribute is present. The first disadvantage can be overcome by using "weighted" generalization of the standard correlation.

B. Community Detection in Attributed Networks

Standard CDAs [8] when dealing with weighted graphs, they fully parse the topological structure of networks, but partially utilize edge attributes accumulated in edge weights, and completely ignore node attribute information.

In social networks edges describe relationships, and their weights strongly depend on roles of participants, which can be represented as node attributes.

CDAs for attributed networks (ANCDAs) are developed for utilizing information available in (1). The presence of different node attributes means heterogeneity of nodes. Also, different edge attributes imply that edges are heterogeneous too. An ideal CDA for attributed networks should provide a balance between structural, and node attributes commonalities and should generate dense clusters with homogeneous vertices' and edges' properties. It is a quite challenging task because these three goals - dense connection as well as edges and nodes homogeneity - can conflict.

Let us briefly review several approaches to CD for attributed networks. In [28] Tian et al. propose graph summarization approach that generates clusters, primarily based on similarity of node attributes, and, at the same time, count edge attributes. Methods introduced in [2,28] represent a group of ANCDAs based on *user-selected node attributes* and combine graph clustering with subspace clustering, where subspace is defined by the selected attributes. Zhou et al. present an SA-Cluster Algorithm [37] and then improve it in an Inc-Cluster Algorithm [38], where new edges are added based on nodes similarity, vertices with identical attributes are connected through additional vertices of node attributes and the Random Walk CDA is applied for constructing augmented graph, where random walk distance matrix is effectively computed by matrix increments. SA-Cluster and Inc-Cluster algorithms are *distance-based* ANCDAs. For such algorithm class, an artificial distance measure combining node attributes and structural information is designed. It uses weights W^1, W^{11} (see Section V) of both these parts, respectively. A drawback of such type of algorithms is that the result of CD highly depends on the parameters W^1, W^{11}, which can not be extracted directly from networks, so, an estimation of this exogenous data might be costly. Another way of the analysis of attributed networks is a *model-based* approach, where a null-model is designed for node and edge information consideration. For instance, Xu et al. in [30] propose a Bayesian Attributed Graph Clustering (BAGC) algorithm, where the following assumptions are used for the null model: a) the true partition exists, but it is unknown; b) vertices from the same community behave similarly whiles nodes from different communities may behave differently. The Bayesian model was used for defining a joint probability distribution, which transforms the attributed network CD problem into a standard probabilistic inference problem that can be solved by a specific variational algorithm. Yang et al. proposed another [35] model-based method CESNA (Communities from Edge Structure and Node Attributes) that in addition to the assumptions a) and b) includes next ones: c) nodes from the same communities most likely are adjacent; d) nodes may belong to multiple communities; e) if more than two nodes belong to the same community, then most likely they are adjacent.

Two of the authors [35], Yang and Leskovec, devoted other research works to the analysis of attributed networks, especially to ANCDAs [32,34]. These researchers validated their algorithms on a number if real networks such as LiveJournal, Friendster, Orkut, Amazon, DBLP with explicit participants characteristics. For instance, in the popular social network Live- Journal it is divisions according to culture, entertainment, life/style, gaming, sports, technology, etc. The real groups of people are considered as ground-truth communities (GTCs), which are used for the validation of different hypothesis and results of CD [32, 34, 35]. It turned out [32] that the GTCs are very different from standard ''structural" communities since CDAs attempt to find tightly connected groups of nodes, which are structural communities, whiles the real GTCs are well separated from each other and not necessarily well connected inside.

A comparison of the sensitivity of numerous scoring functions and how they impact a given community detection algorithm is given [32] along with a CDA based on a local spectral clustering, applying different community scoring functions, and solving SNEP from one seed in each ground-truth community. The results confirmed the hypothesis that for CD of the GTCs is better to use measures of reparability such as conductance.

The authors also state that CDAs for GTCs should allow overlapping the communities since the network participants belong to a various number of GTCs of different categories that typically overlap. The mentioned local spectral clustering algorithm [32] allows detecting overlapping communities by choice of seed nodes in different GTCs. The works [34,35] continue the developing overlapping ANCDAs, but the used approach is entirely different - it is a model-based one. For instance, in [35] there is proposed a graph model that can generate networks with community structure entirely based on the probability of pairs of nodes affiliation to GTCs. The probability serves as a similarity function and is called affiliation function, and the described above CDA CESNA is based on it.

The researchers continued developing the direction in [34] and presented a CDA based on a Cluster Affiliation Model for Big Networks (BIGCLAM). The method uses the same affiliation function and tries to fit nodes to their most likely attribute affiliations according to a model of maximum likelihood when node attribute assignment is approached as an optimization problem.

C. *Analytic Hierarchy Process*

The Analytic Hierarchy Process (AHP) is a multiobjective multi-criteria decision-making method, invented by Saaty [24,25].

An idea of the method is to compare pairwise criteria to obtain relative weights of elements of choice. The challenging part of this method is in assigning weights (global priorities) of alternatives. If alternatives are compared with respect to criteria of the next top level, then these criteria, in turn, are evaluated depending on criteria of the next top level and so on. Weights of all criteria against the next top-level ones form vectors of local priorities, which can be computed accurately for numerical criteria functions or can be assessed based on leading eigenvectors of preference matrices. Prioritisation of decisions is made at the stage of synthesis of local priorities vectors into a global priorities vector, which dimension coincides with ones of local priorities vectors of the decision alternatives. The global priorities vector is a linear combination of these vectors with coefficients depending on local priorities of all criteria and alternatives. Normalization of the local priorities vectors guarantees normalization of the resulting global priorities vector.

V. ACCUMULATION OF NETWORK INFORMATION AND ITS APPLICATION

We utilize available network information in an accumulated network G^{wa} combining a weighted network G^{w} and an association network G^{a}, where G^{w} accumulates structural network data – edges and their attributes; G^{a} combines node attributes.

To be more precise, to form the attributed network G^{wa} we use additional information of three levels. Level I - $\bar{W}^{I} = (W^{I}, W'^{I})$ is a vector of weights of G^{a} and G^{w}, respectively. Level II - $\bar{W}^{II} = (W^{II,k})_{k\in J_K}$ is a tuple of ATNs-weights in G^{a} and $\bar{W}'^{II} = (W'^{II,k})_{k\in J_{K'}}$, is a tuple of ATEs- weights in G^{w}. Level III - $\bar{W}^{III,k} = (W^{III,lk})_{l\in J_{L_k}}$ is a tuple of weights of ATNVs in G^{ak} $(k\in J_K)$ and $\bar{W}'^{III,k} = (W'^{III,lk})_{l\in J_{L_k'}}$ is a tuple of weights of ATEVs in G^{wk} $(k\in J_{K'})$. Notice that $\bar{W}^{I}, \bar{W}^{II}, \bar{W}'^{II}, \bar{W}^{III}, \bar{W}'^{III} > 0$ otherwise dimension of the problem can be reduced and normalized.

We form two sets of auxiliary networks: a) $\bar{G}^{a} = \{G^{ak}\}_{k\in J_K}$ - is a set of association networks corresponding node attributes $\{AT^{nk}\}_{k\in J_K}$; b) $\bar{G}^{w} = \{G^{wk}\}_{k\in J_{K'}}$ - is a set of weighted networks corresponding edge attributes $\{AT^{ek}\}_{k\in J_{K'}}$.

$$||\bar{W}^{I}|| = ||\bar{W}^{II}|| = ||\bar{W}'^{II}|| = \\ ||\bar{W}^{III,k}||_{k\in J_K} = ||\bar{W}'^{III,k'}||_{k'\in J_{K'}} = 1 \tag{9}$$

The weights W^{I}, W'^{I} can be interpreted as priorities of the networks G^{a}, G^{w} in the G^{wa}-network structure. Similarly, $\bar{W}^{II}, \bar{W}'^{II}$ are priorities of the auxiliary networks $\{G^{ak}\}_{k\in J_K}$, $\{G^{wk}\}_{k\in J_{K'}}$. Finally, $\{\bar{W}^{III,k}\}_{k}, \{\bar{W}'^{III,k'}\}_{k'}$ are priorities of subnetworks corresponding to single attribute values. The weights can be obtained in different ways, for instance, by an expert assessment or derived directly from the network.

Remark 2 If information about some of Levels I-III is not available, then we suppose that the corresponding weights in (9) are equal.

A. *Attributed Network Construction*

Node Information Utilization

The One Discrete Node Attribute Network G^{ak} Formation. We build the association network G^{a} as

$$\mathcal{AC}^k = \{ AC_l^k \}_{l \in J_{L_k}} \quad (10)$$

denote a G-partition into ACs related to each value of AT^{nk}.

By (2) each node attribute cluster (AC) with at_l^{nk} - value of AT^{nk} is represented as follows:

$$AC_l^k = \{ v_i \in V : a_i^k = at_l^{nk} \}, | AC_l^k | = n_l^k,$$
$$\sum_{l \in J_{L_k}} n_l^k = n (l \in J_{L_k}, k \in J_K). \quad (11)$$

Let us describe an order of constructing the network $G^{ak} \in \bar{G}^a$. It should be normalized (Condition 1), should have edges between nodes with the same value of AT^{nk} (Condition 2), edge weights within the attribute clusters AC_l^k and $AC_{l'}^k$, $(l \neq l')$ should be proportional to the priorities of the values $at_l^{nk}, at_{l'}^{nk}$ in G^{ak} (Condition 3), and finally, it should have equal weights within the same attribute cluster (Condition 4). In our notations, these conditions can be represented as follows:

- Condition 1 - the normalization (see (7)):

$$\omega(G^{ak}) = 1; \quad (12)$$

- Condition 2 - the edge set formation: $\forall i, j \in J_n$ an edge $\{ v_i, v_j \}$ iff

$$\exists l \in J_{L_k} : a_i^k = a_j^k = at_l^{nk}; \quad (13)$$

- Condition 3 - the edge weights distribution: if (13) holds and

$$\exists l' \in J_{L_k} : a_{i'}^k = a_{j'}^k = at_{l'}^{nk} \Rightarrow \quad (14)$$

$$\frac{w_{ij}^{ak}}{w_{i'j'}^{ak}} = \frac{W^{III,lk}}{W^{III,l'k}} \quad (15)$$

where $\bar{w}^{ak} = (w_{ij}^{ak})_{i,j \in J_n}$ is a WAM of G^{ak};

- Condition 4 - equal weights within AC_l^k:

if for $i, i', j, j' \in J_n$ (13), (14) hold and $l = l' \Rightarrow w_{ij}^{ak} = w_{i'j'}^{ak}$. (16)

Notice that (13) can be rewritten as follows:

$$\exists l \in J_{L_k} : v_i, v_j \in AC_l^k. \quad (17)$$

To satisfy (15), (16) we choose weights $\{ w_{ij}^{ak} \}_{i,j}$ proportionally to $W^{III,lk}$:

$$w_{ij}^{ak} = \nu(k) \cdot W^{III,lk} \text{ if } (17), \text{ otherwise } 0 (i, j \in J_n), \quad (18)$$

where $\nu(k)$ is a normalized factor of G^{ak}.

Define $\nu(k)$ using (11), (12), (18) and the fact that edges exist only between nodes with the same ATNV implying a partition of the network by a disjoint union of complete graphs:

$$1 = \omega(G^{ak}) = \sum_{i,j} w_{ij}^{ak} = \sum_{l \in J_{L_k}} \sum_{i,j \in AC_l^k} w_{ij}^{ak} =$$
$$\sum_{l \in J_{L_k}} \sum_{i,j \in AC_l^k} \nu(k) \cdot W^{III,lk} =$$
$$= \nu(k) \cdot \sum_{l \in J_{L_k}} W^{III,lk} \cdot 2 || K_{n_l^k} || = \nu(k) \cdot$$
$$\sum_{l \in J_{L_k}} W^{III,lk} \cdot n_l^k \cdot (n_l^k - 1), \text{ wherefrom}$$

$$\nu(k) = \left(\sum_{l \in J_{L_k}} W^{III,lk} \cdot n_l^k \cdot (n_l^k - 1) \right)^{-1}, \quad k \in J_K. \quad (19)$$

We summarise the result in the following observation.

Remark 3 Each association network G^{ak} satisfied (12)-(16) is a G^a-partition by complete graphs $\{ K_{n_l^k} \}_{l \in J_{L_k}}$ with the WAM $\bar{w}^{ak}$ defined by (18) and $\nu(k)$ defined by (19), k. We can represent it as:

$$G^{ak} = G^a [\mathcal{AC}^k] = \bigcup_{l \in J_{L_k}} G^a [AC_l^k] =$$
$$= \bigcup_{l \in J_{L_k}} K_{n_l^k}, k \in J_K. \quad (20)$$

Remark 4 Notice that $\bar{W}^{IIIk}$ can be not normalized, but weights (18) do not change if multiply $\bar{W}^{IIIk}$ by a non-zero factor (see (19)). Therefore, without loss of generality, assume that for $\{ \bar{W}^{IIIk} \}_{k \in J_K}$ (9) holds.

Remark 5 If $\bar{W}^{IIIk}$ is unknown, then: a) by Remark 2, weights are equal ($W^{III,lk} = W^{III,l'k}, \forall l, l' \in J_{L_k}$); b) by Remark 4, they are normalized, hence, $W^{III,lk} = \frac{1}{|\bar{W}^{III,k}|} = \frac{1}{L_k}$ and the formulas (19) and (18) become $\nu(k) = \frac{L_k}{\sum_{l \in J_{L_k}} n_l^k \cdot (n_l^k - 1)}, k \in J_{L_k}$,

$$w_{ij}^{ak} = \left(\sum_{l \in J_{L_k}} n_l^k \cdot (n_l^k - 1) \right)^{-1} \text{ if } (17) \text{ holds,}$$

otherwise 0 $(i, j \in J_n)$.

The Discrete Association Network G^a Construction. The association network is formed as a weighted network sum of with the weights:

$$G^a = \sum_{k \in J_k} W^{II,k} . G^{ak} . \qquad (21)$$

Lemma 6 If (12) holds then G^a is normalized:

$$\omega(G^a) = 1 . \qquad (22)$$

Notice that a vertex set of G^a is the same as for the original network: $V(G^a) = V$, its edge set $E(G^a)$ is a union of $\{G^{ak}\}_k$ edges sets: $E(G^a) = \bigcup_{k=1}^{K} E(G^{ak})$, the WAM $\bar{w}^a = (w_{ij}^a)_{i,j}$ is the following linear combination of G^{ak} - are WAMs: $\bar{w}^a = \sum_{k \in J_k} W^{II,k} \bar{w}^{ak}$.

If the node attribute weights $\bar{W}^{II}$ are unknown, then they are supposed to be equal and, similar to $\bar{W}^{III}$ (see Remark 5), we have: $W^{II,k} = \frac{1}{k}, k \in J_k$ and (21) becomes:

$$G^a = \frac{1}{K} \sum_{k \in J_k} G^{ak} .$$

1) Edge Information Utilization

Similar to the association network G^a construction we build the weighted network G^w combining the normalized auxiliary networks $\bar{G}^w = \{G^{wk}\}_{k \in J_{K'}}$ related to the edge attributes $\{AT^{ek}\}_{k \in J_{K'}}$:

$$w(G^{wk}) = 1, k \in J_{K'} . \qquad (23)$$

The networks family $\overline{G}^w$ is built using the same network structure as G:

$$\forall k \in J_{K'}, V(G^{wk}) = V, E(G^{wk}) = E \qquad (24)$$

The One Discrete Edge Attribute Network G^{wk} Formation.

edge attribute AT^{ek}. Similar to (11) for the edge attribute clusters (ECs) we have:

$$EC_l^k = \{\{v_i, v_j\} \in E : a'^k_{\{i,j\}} = at_l^{ek}\},$$
$$|EC_l^k| = m'^k_l (l \in J_{L'_k}, k \in J_{K'}), \qquad (25)$$

where $\sum_{l \in J_{L'_k}} m'^k_l = m, k \in J_{K'}$. Identical to the node partition set $\mathcal{AC}$, we can build $\mathcal{EC} = (\mathcal{EC}^k)_{k \in J_{K'}}$ - a tuple of E-partitions into ECs of different ATEs.

In the same way as the conditions (12)-(16) were used for G^{ak}, each network G^{wk}, $k \in J_{K'}$, satisfies four conditions: a) the normalization condition (23); b) the edge set formation condition (24); c) the weights uniformity within ECs:

$$\text{if for } i, i', j, j' \in J_n \, \exists l \in J_{L'_k} :$$
$$\{v_i, v_j\}, \{v_{i'}, v_{j'}\} \in EC_l^k \Rightarrow w_{ij}^{wk} = w_{i'j'}^{wk} , \qquad (26)$$

and d) the proportion of edge weights within $EC^k, EC^{k'}$ to the weights of the corresponding ATEVs:

$$\frac{w_{ij}^{wk}}{w_{i'j'}^{wk}} = \frac{W'^{III,lk}}{W'^{III,l'k}} \qquad (27)$$

where $\bar{w}^{wk} = (w_{ij}^{wk})_{i,j \in J_n}$ is the WAM of $G^{wk}, k \in J_{K'}$.

Similarly to the network G^{ak}, the conditions (26) are satisfied by choice of the weights $\{w_{ij}^{wk}\}_{i,j}$ proportionally to $W'^{III,lk}$: within EC_l^k $w_{ij}^{wk} = \nu'(k) \cdot W'^{III,lk}$, otherwise 0. Here $\nu'(k)$ is a normalized factor of G^{wk}, which is defined from (23) by (25):

$$1 = \omega(G^{wk}) = \sum_{i,j} w_{ij}^{wk} = \sum_{l \in J_{L'_k}} \sum_{\{v_i, v_j\} \in EC_l^k} w_{ij}^{wk} =$$
$$= \sum_{l \in J_{L'_k}} \sum_{\{v_i, v_j\} \in EC_l^k} \nu'(k) \cdot W'^{III,lk} =$$
$$= \nu'(k) \cdot \sum_{l \in J_{L'_k}} W'^{III,lk} \cdot 2|EC_l^k| = \nu'(k) \cdot \sum_{l \in J_{L'_k}} W'^{III,lk} \cdot 2m'^k_l$$

where

$$\nu'(k) = \frac{1}{2 \sum_{l \in J_{L'_k}} W'^{III,lk} \cdot m'^k_l} , \qquad (28)$$

$$w_{ij}^{wk} = \nu'(k) \cdot W'^{III,lk} \text{ if } \{v_i, v_j\} \in E, \qquad (29)$$

otherwise 0.

Remark 7 Analogically to nodes (see Remark 4), for edges we assume normalization of the weights $\{\bar{W}'^{IIIk}\}_{k \in J_{K'}}$ (see (9)).

Similar to merging nodes with the same values of attributes into attribute

The Weighted Network G^w construction. Similar to G^a construction, we form the weighted network G^w as the weighted network sum of networks $\bar{G}^w$ with parameters $\bar{W}'^{II}$:

$$G^w = \sum_{k \in J_K} W^{II,k} \cdot G^{wk} . \qquad (30)$$

Lemma 8 If (23) holds then G^w is normalized:

$$\omega(G^w) = 1 . \qquad (31)$$

Due to (24), the vertex and edge sets are not changed during the linear network transformation from $\bar{G}^w$ into G^w, hence $V(G^w) = V$, $E(G^w) = E$; a WAM $\bar{w}^w = (w_{ij}^w)_{i,j}$ of G^w is the following linear combination of the WAMs $\{\bar{w}^{wk}\}_{k \in J_{K'}}$: $\bar{w}^w = \sum_{k \in J_{K'}} W'^{II,k} \cdot \bar{w}^{wk}$.

Aggregation of the networks G^a and G^w into the Aggregated Network G^{wa}.

In the current section, we present an approach for the analysis of attributed networks with discrete node attributes and edge attributes. Suppose the association network G^a and the weighted network G^w are formed then the aggregated network G^{wa} is formed as their weighted network sum with coefficients $\overline{W}^I$:

$$G^{wa} = W^I \cdot G^a + W'^I \cdot G^w . \quad (32)$$

Lemma 9 If (23) and(31) hold then G^{wa} is normalized:

$$\omega(G^{wa}) = 1 . \quad (33)$$

Parameters of G^{wa} are the following: a vertex set $V(G^{wa})$ coincides with the set V, an edge set $E(G^{wa})$ is a combination of G^a – and G^w – edge sets, a WAM $\overline{w}^{wa}$ is the corresponding linear combination of the WAMs $\overline{w}^a, \overline{w}^w$. So,

$$V(G^{wa}) = V , \; E(G^{wa}) = E(G^a) \cup E(G^w) ,$$

$$\overline{w}^{wa} = W^I \cdot \overline{w}^a + W'^I \cdot \overline{w}^w .$$

For the final attributed network G^{wa} construction, we perform two steps of aggregating the data of three levels. It is represented by a three-level hierarchy (see Fig. 1).

As it is seen, the networks associated with individual node or edge attributes from the families $\overline{G}^a$ and $\overline{G}^w$ are located on the lowest level. Depending on $\overline{W}^{II}$ and $\overline{W}'^{II}$ these networks constitute different proportions in the next upper level, which consists of the networks G^a and G^w associated with all node and edge attributes. In turn, G^a and G^w form the top level network G^{wa} and participate in the aggregated network to more or less extend depending on W^I, W'^I.

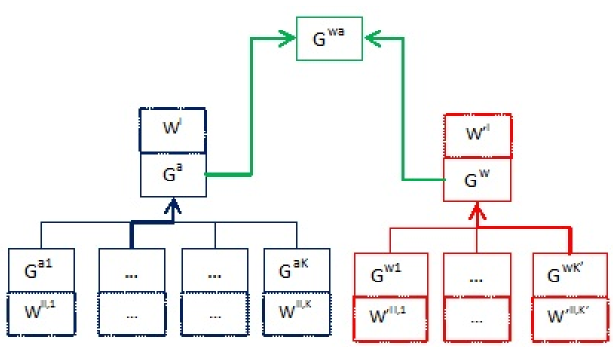


Fig. 1 The aggregated network G^{wa} hierarchy

A generalization of the AHP synthesis phase (see Section 4.c) is based on the observation that instead of vectors of local priorities of alternatives any other objects can be used if their linear combination is defined. For instance, they can be matrices of the same dimension, functions, networks of the same vertex set, etc. The result of the synthesis will be an object of the same type, such as a matrix of a relevant dimension, a function or a network of the same vertex set.

So, at the bottom level of the hierarchy, we use the networks $\{ G^{ak}, G^{wk} \}_k$ of individual node-edge attributes instead of vectors of local priorities of alternatives (see Section 4.c). Then the networks G^a, G^w, and G^{wa} are built as their linear combination. $\overline{W}^{II}, \overline{W}'^{II}$, and $\overline{W}^I$ play a role of vectors of local priorities. This operation is defined since the vertex set is the same for all of the networks. In comparison with the basic AHP, where a numerical global priorities vector is the result of the synthesis phase, the outcome of our procedure is the top level normalized network G^{wa} (see Lemma 9).

B. *Attributed Networks Applications*

In Section V, the weighted network G^w is represented as a linear combination of auxiliary networks of individual edge attributes (see (30)). On the other hand, forming edge set, we assumed that the edges are formed exceptionally based on node attributes. Therefore, CD is conducted on G^w depending on node attributes. Analogically to the associated network G^a representation (21) we represent G^w as a linear combination of auxiliary networks $\overline{G} = \{ G^k \}_k$ related to single node attributes:

$$G^w = \sum_{k \in J_K} W^{II,k} \cdot G^k . \quad (34)$$

Respectively, by (32) G^{wa} can be decomposed into the networks corresponding to ATNs only:

$$G^{wa} = \sum_{k \in J_K} W^{II,k} \cdot G^{wak}, \; where \qquad . \quad (35)$$
$$G^{wak} = W^I \cdot G^{ak} + W'^I \cdot G^k, k \in J_K$$

For such type of networks, our approach to MLCD consists in the following: we run CD on the accumulated network G^{wa}, obtain the initial partition $\mathcal{C}^*$ and define a node attribute AT^{nk_0} underlying the partition. Then we reduce edge weights of G^{wa} by subtracting the network G^{wak_0}, run CD again on the new network $G^{wa} - G^{wak_0}$, obtain a new node partition and define a node attribute underlying it, etc.

C. *Underlying Node Attributes*

Combining the clusters $AC^k_l \in \mathcal{AC}^k$ with the communities in $C^{l^*} \in \mathcal{C}^*$, we obtain $L_k \cdot L^*$ clusters of nodes:

$$AC^k_{ll^*} = AC^k_l \cap C_{l^*}, n^k_{ll^*} = | AC^k_{ll^*} |, l \in J_{L_k}, l^* \in J_{L^*} . \quad (36)$$

Evaluate the following values:

$$z_{ll^*}^{k} = \frac{p_{ll^*}^{k} - p_l^{0k}}{\sqrt{p_l^{0k}(1-p_l^{0k})}}\sqrt{n_{ll^*}^{k}}, l \in J_{L_k}, l^* \in J_{L^*}, \quad (37)$$

where

$$p_l^{0k} = \frac{n_l^k}{n}, l \in J_{L_k}, \quad (38)$$

$$p_{ll^*}^{k} = \frac{n_{ll^*}^{k}}{n_{l^*}}, l \in J_{L_k}, l^* \in J_{L^*} \quad (39)$$

are proportions of the number of nodes in the ACs (11) and the clusters (36) to their cardinality. $\forall C_{l^*} \in \mathcal{C}^*$ underlying node attribute (UNA) of the community corresponds to the maximum value of (37):

$$z_{l^*}^{max} = max_{l \in J_{L_k}, k \in J_K} z_{ll^*}^{k}, l^* \in J_{L^*}. \quad (40)$$

So, $AT^{nk_{l^*}}$ is the community C_{l^*}-UNA if $k_{l^*} \in I_{l^*}$, where I_{l^*} includes indices of ATNs where the maximum (40) is attained:

$$I_{l^*} \subseteq J_K : \forall k \in I_{l^*} \exists l'_k : z_{l'_k l^*}^{k} = z_{l^*}^{max}. \quad (41)$$

Find the number of nodes in the corresponding to (41) clusters:

$$n_k^* = \sum_{k \in I_{l^*}, l^* \in J_{L^*}} n_{l'_k l^*} \quad (42)$$

and take their maximum:

$$k^* : n_{k^*}^* = max_{k \in J_K} n_k^*. \quad (43)$$

AT^{nk^*} is the UNA of the partition.

Note that (40) are One-Proportion test statistics for testing the hypothesis: $H_{ll^*}^{0,k} = \{ p_{ll^*}^{k} = p_l^{0k} \}$ about significance difference of the proportions (39) from the ones (38). If $H_{ll^*}^{0,k}$ is true the statistics (40) is normally distributed ($N(0,1)$). The choice (43) can be justified with the pre-specified level of significance if the null hypothesis $H_{ll^*}^{0,k}$ is considered with an alternate one $H_{ll^*}^{1,k} = \{ p_{ll^*}^{k} > p_l^{0k} \}$. The value at_l^k such that $H_{ll^*}^{0,k} = \{ p_{ll^*}^{k} = p_l^{0k} \}$ is rejected is called *underlying node attribute value* (UNAV) of the community C_{l^*}.

D. The MLCD Algorithm

- **Step 1.** Setup $\tau = 1$ - an initial iteration, $G(1) = G^{wa}$ - an initial network, $I(1) = J_K$;
- **Step 2** Run CD in $G(\tau)$ and obtain the partition $\mathcal{C}^{*\tau}$;
- **Step 3.** Determine $k_\tau \in I(\tau)$ such that AT^{nk_τ} is a UNA of the partition $\mathcal{C}^{*\tau}$ (see 43);
- **Step 4.** Derive G^{wak_τ} from $G(\tau)$: $G(\tau+1) = G(\tau) - G^{wak_\tau}$, $I(\tau+1) = I(\tau) \setminus \{ k_\tau \}$;
- **Step 5**. Assign $\tau = \tau + 1$. If $\tau \le K$ then go to Step 2, otherwise stop.

VI. Experiments

This section is conducted with the help of popular software for Network Analysis - Gephi and IGraph. If a CDA is not specified, then it means that we use the Louvain method [1], which is implemented in both listed programs.

A. Attributed Networks Simulation

Simulate attributed network G^w with each layer formed as a partition by Erdos-Renyi Random Graphs (ERRGs) [7] and the corresponding association network G^a.

Parameters that are common for these networks include: the order $n = 60$, the number of the node attributes $K = 3$, the nodes are divided randomly into $\{ L_k \}_k = \{ 5,4,6 \}$ attribute clusters of the same sizes $(n_l^k)_{l \in J_{L_k}, k \in J_K} = ((n^k)^{L_k})_{k \in J_K} = (12^5, 15^4, 10^6)$, the priorities $\bar{W}^{II} = (0.5, 0.3, 0.2)$ of node attributes.

The auxiliary networks $\bar{G} = \{ G^k \}_k$ are formed as follows: $G^k = G^w[\mathcal{AC}^k] = \bigcup_{l \in J_{L_k}} ERRG(p_l^k, n_l^k), k \in J_K$,

where p_l^k - the probability of an edge $\{ v_i, v_j \}$: $a_i^k = a_j^k = at_l^{nk}$, l, k.

The resulting network G^w is found by (34) and is overlapping of K partitions by ERRGs.

For the corresponding association network, the auxiliary association networks $\bar{G}^a$ are formed as follows:

$$G^{ak} = G^a[\mathcal{AC}^k] = \bigcup_{l \in J_{L_k}} K_{n_l^k}, k \in J_K.$$

The network G^a is overlapping of K partitions by complete graphs, which is defined by (6).

B. Attributed Network Applications

1) MLCD on Synthetic Networks

The effectiveness of the MLCD algorithm (see Section 5.b) is demonstrated with the help of the network G^w generated in Section 6.a for the level of significance *a = 90%* [18,19].

Suppose that G^w is an original network for which we have information about three attributes. The result of CD in G^w is five communities, M = 0.383 which, in fact, do not relate directly to any of the attributes - communities are mixtures of nodes with different attributes. On the other hand, we know that these attributes are the only factors of the network formation, and also we know the weights of each attribute - 0.5, 0.3, 0.2, respectively. The only question here is an identification of these three layers in the network.

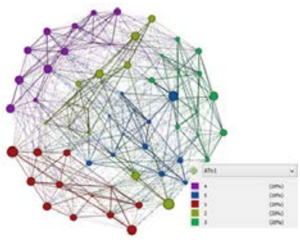


Fig.2 The MLCD

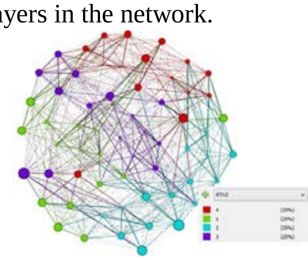


Fig.3 The MLCD result in $G(1)$- $\mathcal{AC}^1 = \mathcal{C}^{*1}$

Fig. 4 The MLCD

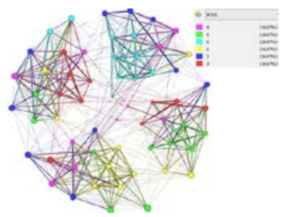


Fig. 5 The MLCD result in $G(2)$- result in $G(2)$, $\mathcal{AC}^2 = \mathcal{C}^{*2}$

According to the MLCD algorithm, we accompany G^w by the corresponding association network G^a (see Section 6.a) and construct an aggregated network (see Section 5.a).

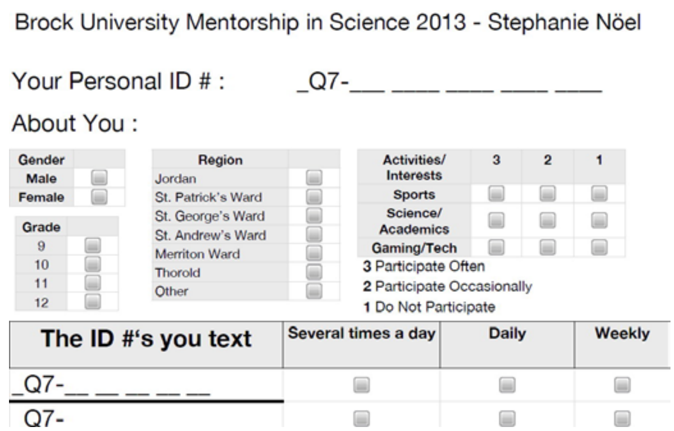
Take equal weights of the structural and attributed parts, $(W^I, W'^I) = (0.5, 0.5)$, where $G^{wa} = 0.5G^a + 0.5G^w$, $\omega(G^a) = \omega(G^w) = 1$. We run the MLCD algorithm on $G(1) = G^{wa}$ and obtain by (42) $(n_k^*)_k = (60,26,21)$ whereby (43) $k^* = 1$. So, $AT^{nk^*} = AT^{n1}$ is the partition $\mathcal{C}^{*1}$-UATN. Moreover, since we obtain the exact partition into attribute clusters related to the first node attribute, $\mathcal{C}^{*1} = \mathcal{C}^* = \mathcal{AC}^1$, AT^{n1} is a dominant node attribute of $G(1)$. An illustration provided in Fig. 2 - 3 shows that attribute clusters related to both AT^{n1} and AT^{n2} can be detected. On the other hand, the weight of the AT^{n1} is $0.5/0.3 = 1.67$ times more than the weight of the AT^{n2}.

Running CD, as expected, we obtain $\mathcal{C}^{*1}$ related to the AT^{n1}. The result of the next, $\tau = 2$, iteration in the network $G(2)$ of the weight $\omega(G(2)) = 1 - W^{II,1} = 0.5$ is presented in Fig. 4. Similar to the previous step we obtain a node partition related to the dominant node attribute AT^{n2} of $\mathcal{C}^{*2}$, $\mathcal{C}^{*2} = \mathcal{AC}^2$, since $(n_2^*, n_3^*)_k = (60,23)$, respectively, $k^* = 2$.

In Fig. 4-5 we can see the division of the network into four communities that are exactly related to the attribute AT^{n2}. The last iteration is conducted in $G(3)$, $\omega(G(3)) = 0.2$, and yields the partition $\mathcal{C}^{*3} = \mathcal{AC}^3$.

2) *The High-School Texting Network (HSTN)*

Also, MLCD was conducted on a real node-edge attributed network, a High-School Texting Network (HSTN) [18,19]. **The HSTN description.** In order to construct G 398 high-school students of Denis Morris Catholic High School (Thorold, Ontario) were asked to provide information about their texting contacts, gender ("Gn"), grades ("Gr"), residence location (Region, "R") and attitude to activities (see Fig. 6).



Brock University Mentorship in Science 2013 - Stephanie Nöel

Your Personal ID # : _Q7-___ ____ ____ ____ ____

About You :

Gender	
Male	☐
Female	☐

Grade	
9	☐
10	☐
11	☐
12	☐

Region	
Jordan	☐
St. Patrick's Ward	☐
St. George's Ward	☐
St. Andrew's Ward	☐
Merriton Ward	☐
Thorold	☐
Other	☐

Activities/ Interests	3	2	1
Sports	☐	☐	☐
Science/ Academics	☐	☐	☐
Gaming/Tech	☐	☐	☐

3 Participate Often
2 Participate Occasionally
1 Do Not Participate

The ID #'s you text	Several times a day	Daily	Weekly
Q7-__ __ __ __ __	☐	☐	☐
Q7-	☐	☐	☐

Fig. 6 The questionnaire

The intensity of: a) texting contacts ("TC"); b) participation in the following activities: Sports ("S"), Science/Academics ("Sc"), and Gaming/Tech ("Ga") is represented by three categories and is ranked from worst to best. For instance, "1" corresponds to the weakest case ("cold contacts and activities participation), "3" - to the strongest one (a "hot" contact and not participation in an activity), and "2" - to medium one ("worm" contacts and rare participation in the corresponding activity).

From each of the completed questionnaires a directed star graph was formed, and then all these graphs were combined into one undirected graph. In case if two participants included each other in their texting contact lists, then such intensity of the mutual contact was established as the highest one.

As a result, the High School Texting Network G $n = 521$ nodes and $m = 1887$ edges was built. The network is decorated by node attributes "Gn","Gr","RL","S","Sc","Ga" and by an edge attribute "TC". The node attribute data is incomplete because from one standpoint, a part of participants did not provide complete information, and from another standpoint, $521 - 398 = 123$ students were mentioned in the contact lists but did not participate in the survey. For these 123 students all the information, except for a partial list of texting contacts and their intensity, is missing.

The HSTN has the following parameters (see Tables 1): a) there are K node attributes (K=6) and K' edge attribute ($K' = 1$); b) node and edge attributes are $\overline{AT}^n = (AT^{nk})_{k \in J_6} = (Gn, Gr, RL, S, Sc, Ga)$ and $\overline{AT}^e = (AT^{e1}) = (TC)$, respectively; c) the node attributes take $\{L_k\}_k$ values $\{L_k\}_k = \{2,3,3,3,7,4\}$ from

$\overline{at}^{n} = (at^{nk})_{k \in J_6} = (\{female, male\}, \{1,2,3\},$
$\{1, 2,3\}, \{1, 2,3\}, \{Jordan,..., Thorold\}, \{9,10,11,12\});$

d) the edge attribute takes 3 values (L1 = 3) from $at^e = at^{e1} = \{1, 2,3\}$. A robust subnetwork G' on n' = 348 nodes and m' = 1148 edges, included those the survey participants who provided all the data, was extracted from G. Table 1 represents $\{nl^k\}_{l,k}$ - frequencies of ATNVs in G'.

MLCD in the HSTN G' Consider the robust HSTN G' and study the nature of communities of its first layer, which are obtained by CD applied to G' directly. For that first, we convert the edge-attributed network G' into a weighted G^w assigning priorities of 1, 2, 3-texting intensity ranks such that to attain maximum modularity among selected combinations: $\bar{W}'^{III}$ = =(0.084,0.275,0.641). Secondly, we assign the node attribute priorities $\bar{W}'^{II}$ = (0.22, 0.14, 0.12, 0.14, 0.13, 0.25) comparing weights of edges within the attribute clusters with the ones between the attribute clusters and the rest of the network. Rest of priorities we take equal according to Remark 2.

l	Gender	n_l^1	S	n_l^2	Sc	n_l^3	Ga	n_l^4
1	female	189	1	77	1	109	1	153
2	male	159	2	77	2	128	2	98
3			3	194	3	111	3	97
Total		348		348		348		348

Table 1 Frequencies of the ATNVs in the HSTN G'

l	Region	n_l^5	Grade	n_l^6
1	Jordan	17	9	90
2	Merriton Ward	60	10	74
3	Other	145	11	103
4	St. Andrew’s Ward	5	12	81
5	St. George’s Ward	4		
6	St. Patrick’s Ward	1		
7	Thorold	116		
Total		348		348

Table 1 Frequencies of the ATNVs in the HSTN G' (cont.)

Since all the node attributes are known, we can compose G'^w with the corresponding association network (see Section 5.a) denoted by $G^{a'}$. Assuming that $\overline{AT}^a$ are only reasons of the attributed network structure formation we choose equal weights of the attributed and structural parts ($W^I = W'^I$ = 0.5) and construct an aggregated attributed network $G^{wa'} = 0.5 G^{a'}$ Ga'+0.5 $G^{w'}$. We conduct CD in the networks $G^{w'}$ and $G^{wa'}$, obtain the node partitions $\mathcal{C}^{*'}$ and $\mathcal{C}^{*'a}$, respectively, and compare the results (see Fig. 7-8):

- the partition $\mathcal{C}^{*'}$ includes $L^{*'}$ = 25 communities. First eight, $\mathcal{C}^{*'(0)} - \mathcal{C}^{*'(7)}$ are the largest ones. They contain at least 2% each of the total number of the nodes and can be seen in Fig. 7. The rest 17 communities - we denote them $\mathcal{C}^{*'(8-24)}$ - primarily correspond to people did not provide their texting contacts, therefore, represented by isolated vertices or by communities with a few nodes. In terms of the HSTN, since only the 8-11-th grades high-school students participated in the survey, it is unlikely that they do not communicate via texting or by another way with someone of classmates. Hence the existence of these isolated components is explained exclusively by lost edges.
- For the network $G^{wa'}$, the situation looks different. The network is one-component, the partition $\mathcal{C}^{*'a,a}$ includes only $L^{*,a}$ = 8 communities, where the first seven contain all nodes except for one. Namely, these communities denoted by $\mathcal{C}^{*'a(0-6)}$ we analyze. Notice that they can be clearly seen in Fig. 8.

These pictures demonstrate an advantage of the usage of network aggregation with an associated network instead of the original network usage. In the aggregated HSTN almost all the mentioned isolated components became connected by the common node attributes edges. As a result, the number of communities decreased significantly and made possible the CD results interpretation. Also, since the number of isolated components decreases a lot it makes possible to restore all the missing network information: we can start with node and edge attributes inference, after that the edge inference of the corresponding associated network can be conducted based on already renovated node attributes [10,18]. In terms of the HSTN, it means that the applying the aggregated networks allows for any node with at least one link or a node attribute to restore the rest of attributes, whiles for the original network it is not possible.

For the HSTN community structure studying the aggregated network provides an opportunity to clarify to which communities the students from the $\mathcal{C}^{*'(8-24)}$ belong if their links to the community were lost during the initial data collecting.

We analyze communities in $\mathcal{C}^{*'a(0-6)}$ and assign underlying ATNVs (UNAVs) to them with the level of significance $\alpha = 0.1$ according to the scheme described in Section 5.b. The result presented in Table 2 is obtained for those four node attributes, for which the communities UNAVs were derived. The other two node attributes, “Science" and “Region", are not significant in the division $\mathcal{C}^{*'}$. In Table 2, the UNAVs for $\alpha = 0.1$ are presented.

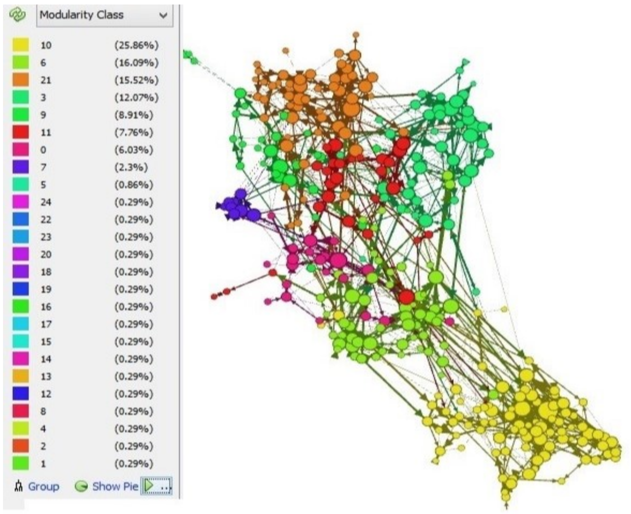


Fig. 7 Communities in the HSTN $G^{w'}$

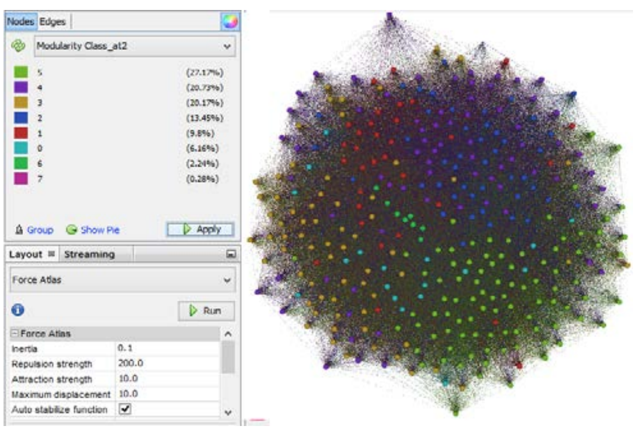


Fig. 8 Communities in the aggregated $G^{wa'}$ (M = 0,314)

It can be seen that only the “Grade"-UNAVs were extracted for each of the communities. Therefore, AT^{n6} = ”Grade” is the dominant node attribute of the partition, as well as for the original G, G' and the accumulated G^{wa}.

For example, review the communities $C_1^*, C_2^*, C_6^* \in \mathcal{C}^{*'a(0-6)}$ of the Grade 11. There are only one “male” community, C_2^*, and two “female” communities: the C_1^* unites girls that like sport and gaming activities, but they are not fans of it; there are mostly girls who definitely do not like gaming in the C_6^*. Two communities relates to the Grade 10, $C_0^*, C_3^* \in \mathcal{C}^{*'a(0-6)}$. For them no underlying attributes among $AT^{n1} - AT^{n6}$ were derived, hence, with the probability 90%, the reason of this division remains undetected. Therefore, the study is not complete.

l*	ni	Gende	Sport	Gamin	Grad
5	95	male			9
4	73		1	1	12
3	71				10
2	47	male			11
1	34	female	2	2	11
0	19				10
6	8	female		1	11

Table 2: $\mathcal{C}^{*'a(0-6)}$-UNAVs, $p = 0.9$

The results of the HSTM analysis show the capacity of detecting minor layers after extracting major ones. Analysis of underlying attributes and values allows studying the nature of the detected communities.

VII. Discussion

CDAs mostly attempt to increase *modularity* [17].

Modularity of the partition $\mathcal{C}^* = \{C_l\}_{l \in J_{L^*}}$ in the unweighted network G is the following characteristics which is maximized:

$$M = \frac{1}{2m} \sum_{l=1}^{L^*} \sum_{u,v \in C_l} \left(a_{u,v} - \frac{d_u d_v}{2m} \right) \to max,; \qquad (44)$$

where d_u is a degree of $u \in V$.

It always can be represented in a form:

$$f(x) = x^T Ax + c^T x \to min, \qquad (45)$$

$$x \in E \subseteq Bn', \qquad (46)$$

$$f_i(x) < 0, i \in J_{m'}, \qquad (47)$$

where $A \in R^{n \times n}$, $A^T = A$, $c \in R^1$, (46) - are direct constraints and (47) are functional constraints. Depending on a choice of E, the problem can be represented with functional constraints or without them.

Various heuristics for solving (44) are known [8, 16,36]. Some of them are implemented in Gephi and IGraph.

However, exact and approximate methods are also can be offered based on the Euclidean statement (45)- (47) of (44). It has a form of an unconstrained quadratic problem over 0 — 1-set:

$$E = \boldsymbol{B}_{L^*}^n(1) = \bigotimes_{i=1}^{n} B_{L^*}(1) \qquad (48)$$

where $B_{L^*}(1)$ is a permutation set induced by $\{0^{L^*-1}, 1\}$. Respectively, (45)-(47) can be treated as a quadratic binary problem with additional linear equality constraints on a sum of a particular set of variables [4,12].

Let us outline original approaches to (44) represented in the form (45),(46),(48) that are based on a vertex locality, a spherical locality, a 2-levelness of $\boldsymbol{B}_{L^*}^n(1)$, its belonging to polypermutation sets [23,40], and other properties of the generalized permutation set [6,22,26]. Its vertex locality allows assuming that the objective function is convex in accordance with [39]. The set is formed as an intersection of a hypersphere with a polytope $\boldsymbol{PB}^n_{L^*}(1) = conv\, \boldsymbol{B}^n_{L^*}(1)$.

That is why the Branch&Bound Polyhedral-Spherical Method (B&B PSM) [21] is applicable for solving (45),(46),(48), as well as approximate schemes of the Polyhedral-Spherical Method (Greedy PSM, GPSM) [21,23].

Another approach is the Functional Representation Method [20], which allows forming a biquadratic continuous functional representation of $\boldsymbol{B}_{L^*}(1)$ [22,23] and then applying Nonlinear Programming to (45),(46),(48).

A peculiarity of $\boldsymbol{B}_{L^*}(1)$ is that a linear function is optimized simply over it, and a projection on the set can be easily found based on its spherical locality [23].

Together with a convexity of the objective function, this allows solve (44) for small dimensional problems and obtain approximate solutions with a prescribed accuracy for higher dimensional ones.

Note that the PSM-group is based on combining two continuous relaxation of the original problem - polyhedral and spherical [40], decomposing $\boldsymbol{B}_{L^*}(1)$ into sets of the same combinatorial type and lower dimensions, and reducing of (45),(46),(48) to similar lower-dimensional problems. Due to the convexity of $f(x)$, the polyhedral relaxation is solved to optimality exactly, and since we deal with the quadratic objective function, the spherical relaxation is solvable exactly as well [39].

Modularity is not only a numerical measure of a presence of communities in networks [8, 33]. Moreover, there is no unique definition of communities [8]. Hence, there is no unique numerical measure *Cr* of community detection quality. The scoring function

Cr can be based on the internal connectivity (Type A) or the external connectivity (Type B); it can also be a combination of the internal and external connectivities (Type C), or it can be based on a network model (Type D) [32]. Thus, if we aim to get a partition with high intra-cluster density, which is the ratio between the number of intra-cluster edges in a community *C* and the number of all possible internal edges of *C*:

$$\delta^{int,C} = \frac{2\,|E(C)|}{|C|(|C|-1)}$$

and low inter-cluster density (7), i.e., the ratio between the number of *C*-inter-cluster edges and the number of all possible *C*-external edges:

$$\delta^{ext,C} = \frac{|E(C,\bar{C})|}{|C|\cdot|\bar{C}|}.$$

of detected communities, then scoring function of C- type may be the following [8]:

$$Cr = \sum_{C\in\mathcal{C}} \left(\delta^{int,C} - \delta^{ext,C}\right) \to max. \quad (49)$$

Taking only one of these goals, we get the criteria of A-type and B-type, respectively:

$$Cr_1 = \sum_{C\in\mathcal{C}} \delta^{int,C} \to max; \quad (50)$$

$$Cr_2 = \sum_{C\in\mathcal{C}} \delta^{ext,C} \to min. \quad (51)$$

The Euclidean statement of (49) is again a binary problem. In case if it's objective function is polynomial, it can be equivalently reformulated in the form (45)-(47), and all the above reasoning works for solving (49).

Note that (49) is a convolution of the two-criterion problem (50), (51), which is also can be reformulated as a two-objective problem (45)-(47),

$$f'(x) = x^T A' x + c'^T x \to max,,$$

where $A' \in R^{n\times n}$, $A'^T = A$, $c' \in R^1$, $A \succeq 0$, $A' \prec 0$, i.e., $f(x)$ is convex, $f'(x)$ is concave. Respectively, multi-objective combinatorial optimization approaches become applicable to resolve our task [11].

VIII. Conclusion

It is clear, we need to develop more effective decomposition into one-layer networks for attributed networks. Such approaches can also be used to attack the problem of restoring missing network data such as lost attributes of edges and nodes. We studied attributed networks, however, not much is known about their formation models, in particular, when the edge weight distribution scheme is considered.

References

[1] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, Fast unfolding of communities in large networks, *Journal of Statistical Mechanics: Theory & Experiment* no. 10, 2008, pp.1-12.

[2] B. Boden, R. Haag, and T. Seidl, "Detecting and exploring clusters in attributed graphs P", in *Conf. on Information & Knowledge Management*, 2013, pp. 2505-2508.

[3] S. Bornholdt and H. G. Schuster, *Handbook of graphs and networks: from the genome to the internet,* 1st ed. Weinheim: Wiley-VCH, 2003.

[4] S. Butenko, P. M. Pardalos, and V. Shylo, Eds., Optimization Methods and Applications: In Honor of Ivan V Sergienkos 80th Birthday, 1st ed. Cham: Springer International Publishing, 2017.

[5] M. Dehmer, Ed., Structural Analysis of Complex Networks, 2011 edition. Dordrecht; New York: Birkhuser, 2010.

[6] G. A. Donec and L. M. Kolechkina, Construction of Hamiltonian paths in graphs of permutation polyhedra, *Cybernet. Systems Anal.,* vol. 46, no. 1, 2010, pp. 7-13.

[7] P. Erdos and A. Renyi, On random graphs, I, *Publicationes Mathematicae (Debrecen*), vol. 6, 1959, pp.290-297.

[8] S. Fortunato, "Community detection in graphs", *Physics Reports,* vol. 486, 2010, pp.75-174.

[9] Ș. Gündüz-Öğüdücü and A. Ș. Etaner-Uyar, Eds., *Social Networks: Analysis and Case Studies,* 2014 edition. New York: Springer, 2014.

[10] E. D. Kolaczyk, *Statistical Analysis of Network Data. New York,* NY: Springer New York, 2009.

[11] L. M. Koliechkina and O. A. Dvirna, Solving Extremum Problems with Linear Fractional Objective Functions on the Combinatorial Configuration of Permutations Under Multicriteriality, *Cybern Syst Anal*, vol. 53, no. 4, 2017, pp. 590-599.

[12] B. Korte and J. Vygen, *Combinatorial Optimization: Theory and Algorithms,* 6th ed. 2018 edition. New York, NY: Springer, 2018.

[13] Z. Lin, X. Zheng, N. Xin, and D. Chen, CK- LPA: Efficient community detection algorithm based on label propagation with community kernel, *PhysicaA: Statistical Mechanics and its Applications,* vol. 416, 2014, pp. 386-399.

[14] B. Liu, Social Network Analysis, in *Web Data Mining*, Springer Berlin Heidelberg, 2011, pp. 269-309.
[15] J. Mcauley and J. Leskovec, Discovering Social Circles in Ego Networks, *ACM Trans. Knowl. Discov. Data,* vol. 8, no. 1, pp. 4:1-4:28, 2014.
[16] R. Missaoui and I. Sarr, Eds., *Social Network Analysis - Community Detection and Evolution,* 2014 edition. New York, NY: Springer, 2015.
[17] M. E. J. Newman, Fast algorithm for detecting community structure in networks, *Physical Review. E, Statistical, Nonlinear, And Soft Matter Physics,* vol. 69, no. 6 Pt 2, 2004, pp. 066133:1066133:5.
[18] O. Pichugina, Community Detection in MultiLayer Networks, Master of Science in Mathematics Thesis, Department of Mathematics and Statistics, Brock University, 2015.
[19] O. Pichugina and B. Farzad, A Human Communication Network Model, in *CEUR Workshop Proceedings* Kyiv, 2016, pp. 33-40.
[20] O. S. Pichugina and S. V. Yakovlev, Continuous Representations and Functional Extensions in Combinatorial Optimization, *Cybern. Syst. Anal.,* vol. 52, no. 6, 2016, pp. 921-930.
[21] O. Pichugina and S. Yakovlev, Continuous Approaches to the Unconstrained Binary Quadratic Problems, in *Mathematical and Computational Approaches in Advancing Modern Science and Engineering,* Springer, Cham, 2016, pp. 689-700.
[22] O. S. Pichugina and S. V. Yakovlev, "Functional and analytic representations of the general permutation set", *Eastern European Journal of Enterprise Technologies,* vol. 79, no. 1(4), 2016, pp. 27-38.
[23] O. Pichugina and S. Yakovlev, "Optimization on polyhedral-spherical sets: Theory and applications", in Conf. Rec. *IEEE 1st Ukraine Conference on Electrical and Computer Engineering, UKR- CON 2017 Proceedings,* Kiev, 2017, pp. 1167-1174.
[24] T. L. Saaty, A scaling method for priorities in hierarchical structures, *Journal of Mathematical Psychology,* vol. 15, no. 3, 1977, pp. 234-281.
[25] T. L. Saaty and J. M. Alexander, *Conflict Resolution: The Analytic Hierachy Approach.* New York: PraegerPub, 1989.
[26] R. Sanyal, F. Sottile, and B. Sturmfels, "Orbitopes" *Mathematika,* vol. 57, no. 2, 2011, pp. 275-314.
[27] J. Scott, Social Network Analysis, Overview of, in *Encyclopedia of Complexity and Systems Science,* R. A. M. Ph.D, Ed. Springer New York, 2009, pp. 8265-8279.
[28] Y. Tian, R. A. Hankins, and J. M. Patel, "Efficient Aggregation for Graph Summarization (Published Conference Proceedings style)," in *Proc. of the 2008 ACM SIGMOD International Conference on Management of Data,* New York, NY, USA, 2008, pp. 567-580.
[29] S. Wasserman and K. Faust, Social Network Analysis: Methods and Applications, 1 edition. Cambridge; New York: Cambridge University Press, 1994.
[30] Z. Xu, Y. Ke, Y. Wang, H. Cheng, and J. Cheng, "A Model-based Approach to Attributed Graph (Published Conference Proceedings style)," in *Proc. of the 2012 ACM SIGMOD International Conference on Management of Data,* New York, NY, USA, 2012, pp. 505-516.
[31] J. Xu, *Topological Structure and Analysis of Interconnection Networks*. Springer Science & Business Media, 2013.
[32] J. Yang and J. Leskovec, "Defining and Evaluating Network Communities Based on Ground- truth (Published Conference Proceedings style)," in *Proc. of the ACM SIGKDD Workshop on Mining Data Semantics,* New York, NY, USA, 2012, pp. 3:1-3:8.
[33] J. Yang and J. Leskovec, "Defining and Evaluating Network Communities Based on Ground truth (Published Conference Proceedings style)," in *Proc. of the ACM SIGKDD Workshop on Mining Data Semantics*, New York, NY, USA, 2012, pp. 3:1-3:8.
[34] J. Yang and J. Leskovec, "Overlapping Community Detection at Scale: A Nonnegative Matrix Factorization Approach (Published Conference Proceedings style)," in *Proc. of the Sixth ACM International Conference on Web Search and Data Mining*, New York, NY, USA, 2013, pp. 587-596.
[35] J. Yang, J. McAuley, and J. Leskovec, "Community Detection in Networks with Node Attributes", in Conf. Rec. *2013 IEEE 13th International Conference on Data Mining,* pp. 1151-1156.
[36] K. Yang, W. Wu, W. Zhang, T. Hsu, and L. Fan, Social Structure Detection, in *Handbook of Combinatorial Optimization*, *P. M. Pardalos, D.- Z. Du, and R. L. Graham, Eds*. Springer New York, 2013, pp. 3089-3132.
[37] Y. Zhou, H. Cheng, and J. X. Yu, Graph Clustering Based on Structural/Attribute Similarities, *Proc. VLDB Endow.,* vol. 2, no. 1, 2009, pp. 718-729.
[38] Y. Zhou, H. Cheng, and J. X. Yu, "Clustering Large Attributed Graphs: An Efficient Incremental Approach", in Conf. Rec. *2010 IEEE 10th Int. Conf. on Data Mining,* pp. 689-698.
[39] S. Yakovlev, Convex Extensions in Combinatorial Optimization and Their Applications, in *Optimization Methods and Applications*, Springer, Cham, 2017, pp. 567-584.
[40] S. V. Yakovlev and O. S. Pichugina, "Properties of Combinatorial Optimization Problems Over Polyhedral-Spherical Sets," *Cybern. Syst. Anal.,* vol. 54, no. 1, pp. 99-109, 2018.

Babak Farzad (Iran, September 7, 1977) Associate Professor of Mathematics, Mathematics and Statistics Department, Brock University, 1812 Sir Isaac Brock Way, St. Catharines, ON L2S 3A1, Canada.

Research interests: Graph Theory, Algorithmic Game Theory, Algorithms and Discrete Optimization.

e-mail: bfarzad@brocku.ca

Oksana Pichugina (Ukraine, September 20, 1966) Associate Professor of Applied Mathematics, Computer Science Department, National Aerospace University "Kharkiv Aviation Institute", 17 Chkalova str., Kharkiv, 61070, Ukraine

Research interests: Combinatorial Optimization, Network Analysis, Graph Theory, Mathematical Modelling

e-mail: oksanapichugina1@gmail.com

Liudmyla Koliechkina (Ukraine, October 22, 1971) Professor of Mathematics, Department of Informatics in Economic Systems, Poltava University of Economics and Trade, 3 Koval St., Poltava, 36014, Ukraine. From 2016, she collaborates with Algorithms and Databases Department, University of Lodz, 3 Uniwersytecka Street, Lodz, 90-137, Poland.

Research interests: Mathematical Modelling, Vector Optimization, Combinatorial Optimization

e-mail: ludaplt1971@gmail.com