þÿModelling Students Thematically Associated Kn
Networked Knowledge from Affinity Statistics

Koponen, Ismo

Springer International Publishing AG
2019

þÿKoponen , I 2019 , Modelling Students Thematically Associated Knowl
Knowledge from Affinity Statistics . in S P Cornelius , C Granell Martorell , J
Gomez-Gardenes & B Gonçalves (eds) , Complex Networks X : Proceedings of the 10th
Conference on Complex Networks CompleNet 2019 . Springer Proceedings in Complexity ,
Springer International Publishing AG , Cham , pp. 123-134 , Conference on Complex
Networks CompleNet 2019 , Taragona , Spain , 18/03/2019 . https://doi.org/10.1007/978-3-030-14459-3_10

http://hdl.handle.net/10138/310189
https://doi.org/10.1007/978-3-030-14459-3_10

unspecified
acceptedVersion

Modelling Students' Thematically Associated Knowledge: Networked Knowledge from Affinity Statistics

Ismo T. Koponen

Abstract Students' knowledge is often organized around relations and key concepts but it sometimes also resembles associative knowledge, where connections between knowledge elements are based on thematic resemblance without overarching organization based on substantiation or logical reasoning. Because it is known that associative knowledge, while important for learning too, may be very differently structured from more organized knowledge, a closer look on students' thematically associated knowledge is warranted. In this study we model students' thematically associative knowledge as a network of pairwise associative connections. The model is based on the assumption that associative knowledge is by a large degree governed by the intrinsic affinity of the knowledge elements that consists of the thematically associated knowledge base. The model introduced here makes minimal assumptions about the affinity distribution of such knowledge. The results show that in this case, under very general conditions, the network of associative knowledge is characterized by inverse power laws of degree, eigenvector, and betweenness centralities. These results agree with the empirically found properties of students' associative networks.

1 Introduction

Knowledge acquisition and processing strategies are specific for the context of learning and for type of targeted knowledge. A starting point for learning is often familiarization with key concepts or key knowledge items, which then are processed further and integrated into more coherent knowledge structures [9, 15]. In that knowledge processing, students' familiarization with the target knowledge often starts with proposing thematic connections between the knowledge items and their possible relationships, for example in the form of concept maps or mind maps.

Such connections can be taken as basically pairwise (dyadic) thematic word or term associations, where associative connections are established on the basis of thematic resemblance or kinds of family resemblance, but where no detailed substantiation or justification of connections is provided [17, 19]. The knowledge processing may then continue with more organized strategies and concept map-type scaffoldings in the form of integrated knowledge systems [9], known to be useful in a variety of learning contexts, for example equally useful in learning science [1, 11, 15, 18, 19] as in learning history [22].

Cognitively oriented research of learning claims that associative knowledge is different from knowledge which is structured through more complex relational dependencies [7, 8, 15]. This may well hold also for thematically associated knowledge, where common theme or topic is the basis of associative connections. The difference between such thematically associative and structured knowledge was noted also in a study focusing on how students organize their knowledge using concept maps [14]. A recent empirical study [13] shows that in the context of history of science, at least students' preliminary knowledge is structured differently than substantiated knowledge [12]. To understand better students' knowledge organization strategies a recent study modelled it in the form of concept networks by using simple linkage-motifs to generate the concept networks. The model produced networks which structurally closely match concept networks made by students, when the construction of networks is rule-based [12]. In that case, networks have degree distributions which are centered and thus have a scale, and have a relatively high clustering [12]. Motivated by the notion that thematically associative knowledge[13] may be very different from rule-based, relational knowledge [12], we focus here on modelling associative networks, which are based on pairwise (dyadic) knowledge item associations and model such networks as affinity-based networks.

We model here knowledge item networks and their properties as they are reported in a recent study addressing the thematically associative knowledge items in the context of history of science [13]. In that study a group of 25 students was involved. The set of knowledge items came from preparatory tasks to explore the history of science over 3 centuries between 1550 and 1850 and how that history of science was embedded as a part of the culture, society, and politics of that same era [13]. This data consisted of about 1300 different knowledge items and about 2500 different pairwise thematically associative connection between them. The study showed that resulting network was characterized by a fat-tailed distribution of node degrees [13]. Also betweenness centrality was found to be distributed according to fat-tailed distribution. In all cases, the distributions were reasonably well-fitted by an inverse power law distributions, which allowed to describe them by using a single relevant parameter, power λ, where $1 < \lambda < 2$. In what follows, we refer to such heavy-tailed distributions as inverse power laws. The present study concentrates on rationalizing a generative model, which can reproduce the inverse power laws of centrality distributions as found in the real thematically associative networks.

We show that to reproduce the empirically found properties of students' networks as reported in Ref. [13] we can construct a kind of a minimal model as an affinity-

ismo.koponen@helsinki.fi

based network [2, 6, 21], with very minimal assumptions about the distribution of the affinities. The resulting network comes close to the empirically found networks when students produce connection between knowledge elements based on thematic associations [13], but are very different from networks found in cases where students' organize their knowledge in rule-based ways [12]. The results support the notion that the ways the students handle knowledge organized around associative connections leads to very different knowledge organization in comparison to situations, where they use relational, rule-based dependencies; both strategies lead to simple regularities but different regularities.

2 The Empirical Case: Associative Knowledge

The empirical findings of associative network and its properties that will be modelled here are based on results recently reported in an empirical study addressing how students make thematic associative connection between different knowledge items in a science (physics) history course for a third and fourth year students (pre-service teachers). The course aim was to discuss the history of physics as part of science history, part of the history of humanities and arts, and as part of general history, in expanding circles. The results reported in Ref. [13] are based on data coming from pre-tasks on which students explored and constructed connections between the historical characters, scientists, ideas, inventions and institutions etc. they thought were of major interest or importance for history of science and history in general. Students reported the connections in form of pairwise associations (dyads), for example [`galilei` $\leftrightarrow$ `heliocentricmodel`]. That data was used in Ref. [13] to construct a complex, thematically associative network of about 1300 nodes and 2500 links. Further details of the course, analysis of the empirical and results are reported in Ref. [13].

The main finding of the empirical study was that thematically associative networks, in the group-level when all student networks were collated, had heavy-tailed distribution of degree centrality D and betweenness centrality B. What is of interest here for modelling is the result that values of D and B turned out to be heavy-tailed and to have approximately an inverse power law type distribution with the inverse power $\lambda \in [1.5, 2.0]$. Of course, the networks were not scale invariant and inverse power law should be taken only as an appropriate fit and in sense revealing the heavy-tailed nature of distribution of values D and B [13].

3 The Model

The basic assumption of the model is that the structure of the students' thematically associative knowledge as it is captured by the network consisting of all different pairwise connections is determined solely by the intrinsic affinity α_k of the knowledge elements $k = 1, 2, \ldots, N$. The intrinsic affinity is for some knowledge elements substantially higher than for some other elements. The formation of links between knowledge elements is determined also solely by their affinity. The affinity α_k, however, cannot be directly available to students. A more plausible assumption is that affinity related ranking R_k of knowledge elements is the basis for forming the linkages. Here, we assume that the appropriate ranking is simply equal to the cumulative distribution R_k of affinities

$$R_k = \sum_i^k \alpha_i / \sum_i^N \alpha_i, \quad R_k \in [0, 1] \tag{1}$$

We next assume that a characteristic value $\bar{R}$ exists, which may depend on task, time allowed for the task, and the average competency of students participating in completion of the task. The probability π_k that a given knowledge element k is linked to another knowledge element is then assumed to correspond to a maximally uncertain choice under this simple constraint. The probability of formation of a link between knowledge elements p and q is then assumed to be proportional to the product $\pi_p \pi_q$.

The probability π_k is now through maximization of the information theoretical (Shannon-Jaynes) entropy function [10]. Here, in what follows, to allow as broad a generality as possible, we adopt the generalized (Tsallis) q-entropy [16, 23–25] in the form

$$I_q = \frac{1}{q}\left[1 - \sum_i \pi_i^{1+q}\right], \quad q \in \,]-1, 1[\tag{2}$$

The exponent q governs the non-extensivity of the entropy. The normal, extensive Shannon-Jaynes entropy $I = -\sum_i \pi_i \log \pi_i$ is recovered at the limit $q \to 0$ [16, 23]. The next step is then to introduce multipliers for variational maximization of the entropy function in Eq. (2). The resulting distribution π_k, which maximizes the q-entropy given the constraint $\bar{R}$ = constant, is a q-exponential [16, 25] (for details of derivation, see Ref. [16])

$$\pi_k = \pi_0 \exp_q[-\beta R_k], \quad \text{where} \;\; \exp_q[x] = \left[1 + \frac{q}{1+q}\, x\right]^{1/q} \tag{3}$$

The function $\exp_q[x]$ is a q-deformed (or q-generalized) exponential function which is reduced to the normal exponential function in limit $q \to 0$. The parameter β is the multiplier corresponding to the constraint that $\bar{R}$ is kept constant. Note that now $\beta < 0$ because $R_k \to 1$ when $k \to \infty$ and $\exp_q[x]$ must be an increasing function [16, 25]. Another multiplier corresponding to the normalization condition is absorbed in normalization π_0. As the functional form of π_k is now known, we require that $\pi_k \to 1$, when $R_k \to 1$. This fixes the normalization coefficient to a value $\pi_0 = 1/[1 - \beta q/(1+q)]$, where $\beta < 0$.

A similar result as in Eq. (3) is obtained through entirely different chain of arguments, starting from affinity distribution and finding a linking probability, which leads to an inverse power law type degree distribution but which does not directly depend on affinity distribution but only through the cumulative distribution R_k [21]. The derivation in Ref. [21] shows that it is always possible to find an affinity distribution which satisfies Eq. (3) for a given power λ of the inverse power law for degree distribution for node degrees d of the form

$$P(d) \propto d^{-\lambda}, \quad \lambda \geq 1 \tag{4}$$

The advantage of derivation in Ref. [21] is explicit connection between the parameters appearing in linking probability to the power of degree distribution in Eq. (4) and to minimum and maximum degrees $k_{\min}$ and $k_{\max}$, respectively, allowed by the choice of the parameters. We utilize these results and rewrite the parameter dependencies of Eq. (3) as follows:

$$q = 1 - \lambda, \quad \lambda \in \,]1, 2[\tag{5}$$

$$\beta = \frac{1 - r^{\lambda - 1}}{1 - \lambda} < 0 \quad \text{and} \quad \pi_0 = \frac{r^{3-\lambda}}{1 - r^{2-\lambda}} \tag{6}$$

where $r = k_{\min}/k_{\max} \ll 1$. With this parameterization the linking probabilities π_k should lead to an inverse power law degree distribution $P(d)$ with power λ. In what follows, we generate networks based on the linking probability given by Eq. (3) and with parameters λ and r as defined by Eqs. (4)–(6).

4 Simulations

The simulations to generate networks and their analysis are carried out by using the `IGraph` package [3]. `IGraph` provides functionality for generating efficiently affinity-based networks simply by providing the probabilities π_k for the routine `IGStaticFittnessGame`. The output of the routine is network with a pre-determined number of links, linked according to the probabilities π_k. From these networks, we measure distributions of: (1) Degree centrality D, (2) eigenvector centrality E, (3) betweenness centrality B, (4) Katz-centrality K, and (5) closeness centrality C_C. In addition, the average value $\bar{C}_L$ of local clustering coefficient and assortativity A is obtained. The definition of these standard network measures is as usual [4, 5]. As a null-model to compare the stability of results when linkages are changed we use networks obtained by rewiring the simulated network but preserving the degree sequence. In all rewirings, the `IGraph` routine `IGRewire` is used with 15,000 rewirings. First, we study the minimal model where no modular structure is introduced. Second, effects of modularity are studied by using a simple, stratified model of modularity. Modules are introduced in three levels $L = 1, 2$, and 3, so that level L has 3^{L-1} modules, where $M_L = \alpha^{L-1}M_0/(9 + 3\alpha + \alpha^2$ with $\alpha \in [1, 3]$ is the number of links in each set of module, not contained in lower level modules. M_0 is the total number of links. Initially, a set of affinities is assigned to all 9 level 3 modules. As will be seen, in case of present minimal model rewired networks are expected to have identical properties in comparison to the original ones, while modular structure affects slightly the eigenvector, betweenness, and Katz-centralities.

Simulations are carried out for networks of sizes $N = 300, 600$, and 1200, corresponding to sizes from the smallest to the most extensive one found in the empirical samples and in the aggregated sample. The values of lambda studied are $\lambda = 1.1$, 1.3, 1.5, and 1.9, also corresponding to the lowest and highest values found in the empirical sample. Similarly, values of the parameter r are chosen from $r = 0.5 \times 10^{-1}$ to 8×10^{-3} to correspond to a cut-off degree in the range from 100 to 200, slightly more than expected in the empirical case where degree 70 is maximum. With these choices, the number of nodes connected in simulations is comparable to empirical cases. No detailed match, however, with empirical data is attempted, because the data-sets are too limited to allow a meaningful quantitative comparisons. The simulations and parameters are chosen to provide only a plausible qualitative agreement with the empirical data.

5 Results

Simulation results for the distributions of degree, eigenvector, and betweenness centralities D, E, and B, respectively, are shown in Fig. 1. The results show that these distributions can be reasonably well fitted with the inverse power law as predicted. The detailed value of the powers of inverse power law fits slightly depends on parameter r. In Fig. 1 the average corresponding three different choices of r are reported. The values for the powers λ_X for centralities $X \in \{D, E, B\}$ are obtained by fitting the inverse power law to the simulated distributions shown in Fig. 1. The detailed breakdown of simulation results for the powers λ_X obtained from simulations with $\lambda = 1.3, 1.5, 1.7$, and 1.9 and for three different values r and the corresponding averages are reported in Table 1. The Katz-centrality is not shown for these cases, because it cannot be fitted with power law very reliably and, moreover, for the simple, un-modular model provides no additional information. In modular case, however, it becomes useful and provides information of the effects of modularity. The average values $\bar{C}_C$ and $\bar{C}_L$ of closeness centrality and local clustering coefficient, respectively, and the assortativity A are also reported. In all cases, the robustness of results under rewiring which preserved the degree sequences was tested by using the rewiring `IGRewire`. The slight changes in distributions that were observed are too small to be distinguishable in the scatter plots in Fig. 1.

The power λ_D obtained for the degree centrality distributions is in all cases slightly larger than the parameter value λ which, according to the theoretical

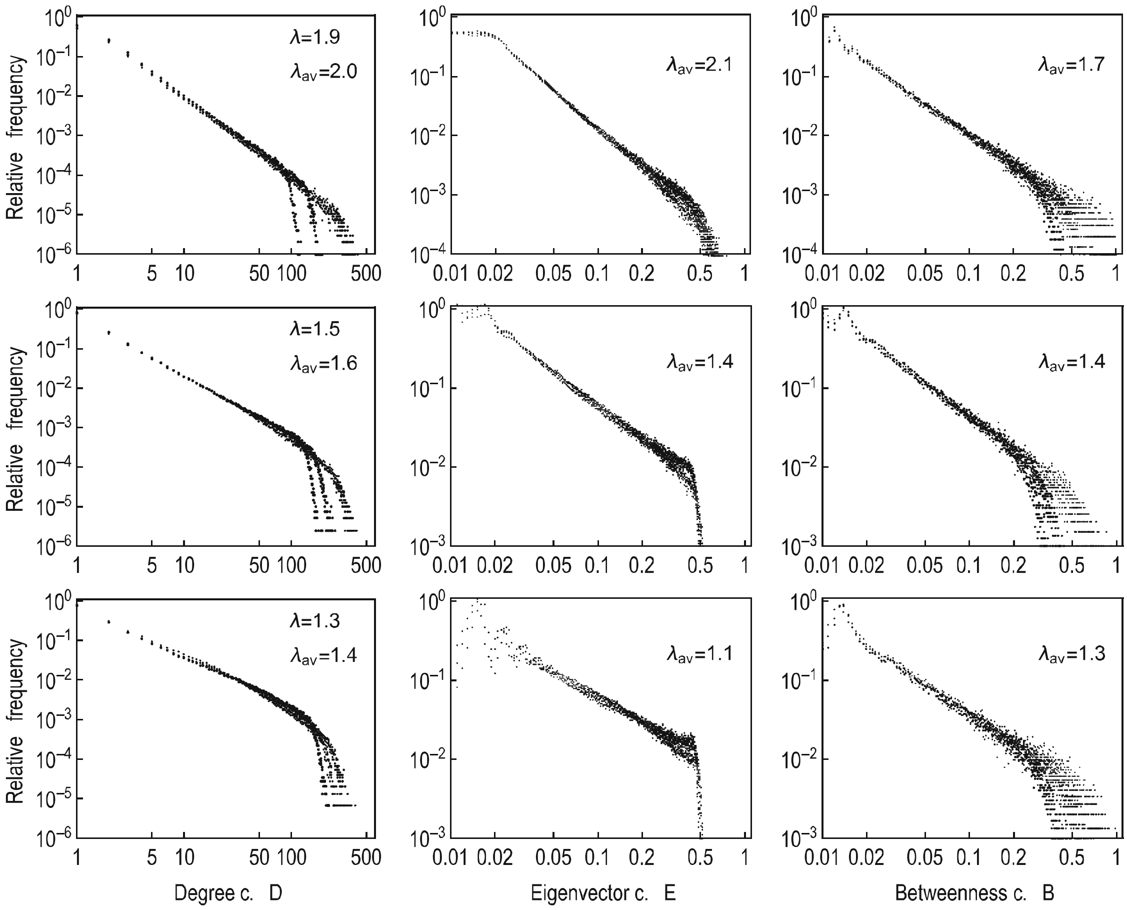

Fig. 1 Distributions of degree (D) , eigenvector (E), and betweenness (B) centralities. The values of E and B are scaled to range from 0 to 1. Results are shown for parameter $\lambda = 1.3, 1.5$, and 1.9, in each case for three different cut-off parameters β. In each case λ_{av} is the average of the best fits to inverse power law -type part of curves as reported in Table 1. All results are based on 1000 repetitions

prediction, should determine the power of degree distribution and $\lambda_D = \lambda$. However, with increasing values of λ and decreasing values of r, the cut-off effects and finite size effects are reduced, inverse power law dependence spans a more extensive region of data and, consequently, $\lambda_D \rightarrow \lambda$.

The eigenvector and betweenness centralities reasonably well follow power laws when the values of these centralities are large enough and exceed the relative value of 0.02. The power law dependence is expected on the basis of previous studies, where affinity-based models with power law degree distribution are also seen to have inverse power law distribution of betweenness centralities [2, 6]. The inverse power law distribution of eigenvector centralities, on the other hand, is expected based on the fact that betweenness and eigenvector centralities also often have substantially high correlation [20]. The finding that degree, betweenness, and eigenvector centralities all follow the inverse power law distribution is also in agreement with empirical results, where similar behavior is observed [13]. The results of the simulations for D, E, and B support an interpretation that the

Table 1 The powers λ_X for centralities $X \in D, E, B$ obtained from simulations for $\lambda = 1.3, 1.5, 1.7$, and 1.9

λ	r $\times 10^{-3}$	λ_D ± 0.01	λ_E ± 0.01	λ_B ± 0.03	$\bar{C}_C$ ± 0.05	$\bar{C}_L$ ± 0.05	A
1.3	2	1.56	1.28	1.25	0.18	0.62	−0.58
$N = 300$	5	1.43	1.03	1.26	0.16	0.60	−0.55
	8	1.30	0.86	1.26	0.16	0.56	−0.50
	AV	1.4	1.1	1.3	0.17		
1.5	2	1.62	1.52	1.46	0.27	0.41	−0.47
$N = 700$	5	1.56	1.33	1.42	0.25	0.31	−0.36
	8	1.49	1.33	1.41	0.24	0.24	−0.27
	AV	1.6	1.4	1.4	0.25		
1.7	0.5	1.94	2.08	1.64	0.33	0.31	−0.39
$N = 900$	2	1.82	1.73	1.60	0.30	0.21	−0.32
	5	1.75	1.73	1.54	0.28	0.12	−0.19
	AV	1.8	1.8	1.6	0.30		
1.9	0.5	2.00	2.26	1.78	0.37	0.13	−0.25
$N = 1200$	2	1.97	2.00	1.71	0.33	0.07	−0.16
	5	1.93	2.01	1.70	0.32	0.04	−0.09
	AV	2.0	2.1	1.7	0.34		

The average values $\bar{C}_C$ and $\bar{C}_L$ of closeness centrality and local clustering coefficient, respectively, and the assortativity A are also reported. The average values over all r are denoted by AV and given with two significant digits

properties of the associative networks can be understood as a consequence of the simple affinity based linking, taking place under conditions of very minimal information of the absolute affinities; what is needed is only the ranking of affinities and a constraint determining the average (or any other single characteristic value) characterizing the rankings.

The closeness centralities obtained in the simulations also show the variation under the changes in λ and r, similarly also the average value of local clustering. In addition, assortativity also changes substantially when the parameter λ is changed from 1.3 to 1.9. This is contrary to findings in empirical cases, where all these values appear to be rather constant [13]. The statistical uncertainties of the empirical sample are considerable and thus also this conclusion is tentative. We believe that the closeness centrality, clustering, and assortativity are sensitive to the modular structure of the networks. The empirical thematically associative networks are modular, although the modularity is not striking, there are clear signs of it. The community search based on modularity optimization by `IGCommunitiesOptimalModularity` finds about from 15 to 20 communities. Here, however, we will not pursue this question further since we are focusing here only on the question if simple model based on affinity is enough to explain the emergence of power law distribution for D, E, and B in case of associative networks.

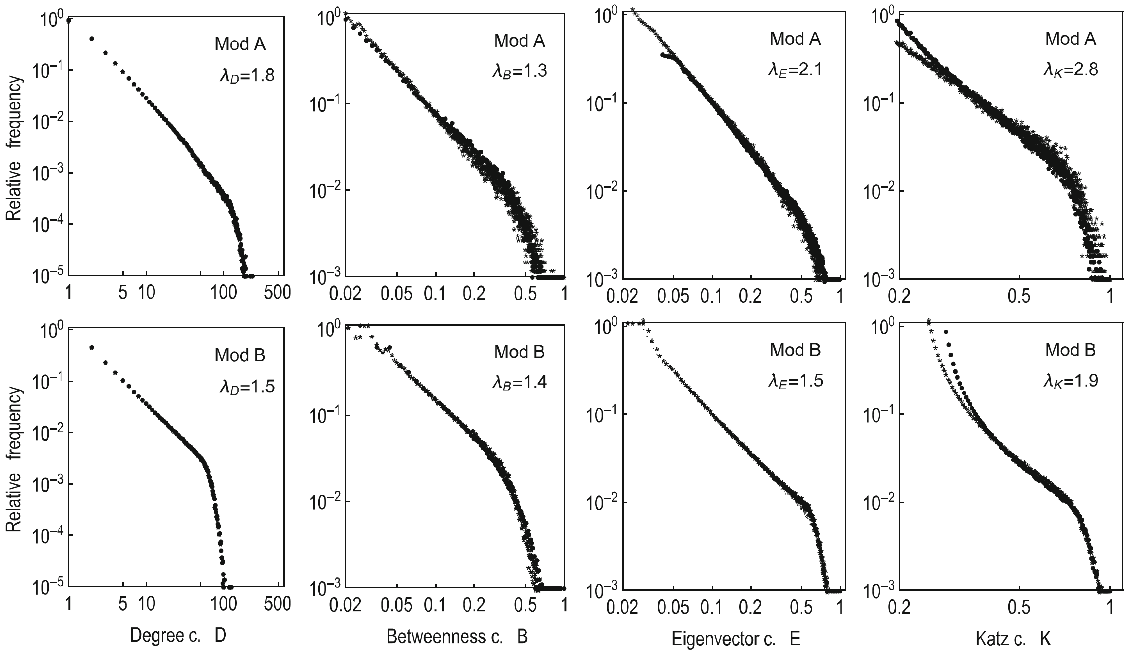

Fig. 2 Distributions of degree (D) , eigenvector (E), betweenness (B), and Katz (K) centralities for modular networks. All values of E, B, and K are scaled to range from 0 to 1. Results are shown for parameter $\lambda = 1.5$ with $r = 5 \times 10^{-3}$. and for modularity with $\alpha = 1$, $N = 1000$ and $M = 4000$ (Mod A) and $\alpha = 3$, $N = 1200$ and $M = 5000$ (Mod B). The results for modular networks (bullets) and rewired counterparts (stars) are shown in same plots (but for most parts, are not distinguishable). The average of the best fits to inverse power law -type part of curves for modular networks is as reported in Table 2. All results are based on 1000 repetitions

The effect of modularity on the distributions of centralities is not large, but is nevertheless detectable. The effect of modularity was studied in detail for $\lambda = 1.5$ and for different modularities with the parameters $\alpha = 1$ with $N = 1000$ and $M = 4000$ (Modularity A); and $\alpha = 3$ with $N = 1200$ and $M = 5000$ (Modularity B), where $\alpha = 1$ corresponds case where the largest module contain the same number of links as the smallest modules, while for $\alpha = 3$ the largest module has equally many links as all small modules. The effect of modularity disappears with rewiring and thus rewired networks provide the benchmark results to quantify the effect of modularity. Figure 2 shows the degree (D), betweenness (B), eigenvector (E), and Katz (K) -centralities for network with $\lambda = 1.5$ and $\alpha = 1$ and 3. The corresponding values of powers λ_X, $X \in \{D, B, E, K\}$ are provided in Table 2. As is seen, the effects of modularity, when results for modular networks (mod) are compared with their rewired counterparts (rwd), are rather small, except for Katz-centrality. Increasing modularity (i.e., decreasing the value of α from 3 to 1) increases the values of powers λ_D and λ_E but decreases λ_B. This signals that with increasing modularity local connectivity increases, but nodes with high values of betweenness centrality remain and their role in connecting the modules becomes more important in comparison to other nodes. The effect is weak, but this tendency can be made more visible by introducing divergence $\Delta[B, E] = (\delta[B] - \delta[E])$, where $\delta[X] = \lambda_X - \lambda_X^{\mathrm{rwd}}/(\lambda_X + \lambda_X^{\mathrm{rwd}})$ quantifies the difference of power λ in

Table 2 The powers λ_X for centralities $X \in \{D, E, B, K\}$ obtained from simulations for $\lambda = 1.5$ and $r = 5 \times 10^{-3}$

N /M_0	α	λ_D ±0.01	λ_E ±0.01	λ_B ±0.03	λ_K ±0.05	Δ	$\bar{C}_C$ ±0.05	$\bar{C}_L$	A
Mod A	1, mod	1.79	2.09	1.33	2.82	0.33	0.32	0.45	−0.35
	1, rwd	1.79	1.81	1.60	2.22		0.18	0.17	−0.25
1000/	2, mod	1.58	1.62	1.34	2.37	0.12	0.24	0.25	−0.27
4000	2, rwd	1.58	1.56	1.46	2.20		0.13	0.13	−0.20
	3, mod	1.51	1.49	1.42	2.27	0.05	0.14	0.15	−0.20
	3, rwd	1.51	1.45	1.45	2.20		0.12	0.10	−0.16
	1, mod	1.86	2.04	1.39	2.81	0.24	0.38	0.46	−0.40
	1, rwd	1.85	1.93	1.68	2.21		0.12	0.12	−0.18
1200/	2 mod	1.60	1.66	1.40	1.95	0.16	0.31	0.27	−0.28
5000	2, rwd	1.60	1.55	1.52	2.26		0.15	0.14	−0.22
	3 mod	1.53	1.49	1.44	1.90	0.04	0.16	0.17	−0.20
Mod B	3, rwd	1.51	1.44	1.45	2.03		0.12	0.10	−0.17

The modularity is imposed by using $\alpha = 1, 2$, and 3. The most sensitive measure for effect of modularity is the divergence $\Delta \equiv \Delta[B, E]$. Results are given for modular network (mod) and rewired network (rwd). The average values $\bar{C}_C$ and $\bar{C}_L$ of closeness centrality and local clustering coefficient, respectively, and the assortativity A are also reported. Results for Mod A and Mod B are shown in Fig. 2

non-rewired network to corresponding power λ^{rwd} obtained for rewired network. With rewiring, when modularity decreases (i.e., α increases), $\Delta[E, B]$ approaches very low values indicating that effects of modularity diminish. This tendency is also manifest in behavior of Katz-centrality. Although the distribution of values of K cannot be reliably fitted with an inverse power law, such a fit reveals a clear tendency which shows that relative frequencies of high K-values decrease with increasing modularity. This supports the interpretation that with increasing modularity the role of long-paths diminishes in comparison to local connectivities and shorter paths.

Simulations performed for other choices of parameters λ, N and M_0 and α provide essentially similar trends as shown by distributions in Fig. 2 and corresponding powers of power law fits summarized in Table 2. The results suggest that the effects of modularity remain very weak for affinity distribution as given by Eq. (3) and parameters corresponding inverse power law distributions.

6 Discussion and Conclusions

We have explored the possibility that students' thematically associative knowledge can be understood as affinity-based process of making associative connections. The motivation for the study stems from the notion that when students' associative knowledge of science history is arranged in the form of a network, consisting of linked pairwise thematically associative connections, a network with inverse power law distribution of degree and betweenness centralities emerges [13].

The basic hypothesis of the study is that the inverse power laws are signatures that thematical associations are based on simple statistics, which is robust enough to nearly invariably lead to inverse power law type dependence. Starting from this hypothesis, we derived the linking probability by assuming that it on the relative rankings of affinities instead of requiring detailed knowledge of the distribution of affinities. In this case, the linking probabilities are obtained as an outcome of maximization of information theoretic entropy, i.e. as outcomes of maximal uncertainty.

The linking probability distribution thus derived was used to simulate the networks. For simulations, `IGraph` and its routine `IGStaticFitnessGame` were adopted. The degree, eigenvector, and betweenness centralities of simulated networks matched closely enough the empirical properties of students' associative networks to support the hypothesis put forward in deriving the distribution. The hypothesis that for major part the properties of the real thematically associated networks may depend only on the affinity rankings and not on the observable modularity of such networks was tested by imposing stratified modularity on the model networks. It was observed that although the modularity affects the centrality distributions, the effect is generally very weak. Therefore, the results support the view that students' associative knowledge (of science history) may be indeed governed by their overall conceptions of the affinity or importance of the knowledge elements (characters, ideas, inventions and events, etc.) forming their knowledge base.

The structure of thematically associative networks is clearly different from the structure of concept networks based on detailed knowledge substantiation [12, 13]. In the case of substantiated knowledge networks, the degree centrality distribution is peaked, resembling a gamma-distribution[12] rather than an inverse power laws as in case of associative knowledge [13]. This notion supports the claim that students' thematically associative and relationally structured knowledge is differently organized, and may actually represent different forms of knowledge [7, 8, 14]. The results of the present study and the study related to organization of substantiated knowledge [12] are, of course, still far from justifying such a conclusion but provide nevertheless promising network based methods to test such arguments.

Acknowledgements This research was funded by the Academy of Finland, Grant 311449.

References

1. Amadieu, F., van Gog, T., Paas, F., Tricot, A., Marine, C.: Effects of prior knowledge and concept-map structure on disorientation, cognitive load and learning. Learn. Instr. **19**, 376–386 (2009)
2. Caldarelli, G., Capocci, A., Rios, P.D.L., Munoz, M.: Scale-free networks from varying vertex intrinsic fitness. Phys. Rev. Lett. **89**, 258702 (2002)
3. Csárdi, G., Nepusz, T.: The igraph software package for complex network research. InterJ. Complex Syst. **1695**(5), 1–9 (2006)

4. da Costa, L.F., Rodrigues, F.A., Travieso, G., Boas, P.R.V.: Characterization of complex networks: A survey of measurements. Adv. Phys. **56**, 167–242 (2007)
5. Estrada, E.: The Structure of Complex Networks. Oxford University Press, Oxford (2012)
6. Goh, K.I., Kahng, B., Kim, D.: Universal behavior of load distribution in scale-free networks. Phys. Rev. Lett. **87**, 278701 (2001)
7. Goldwater, M.B., Schalk, L.: Relational categories as a bridge between cognitive and educational research. Psychol. Bull. **142**, 729–757 (2014)
8. Halford, G.S., Wilson, W.H., Phillips, S.: Relational knowledge: the foundation of higher cognition. Trends Cogn. Sci. **14**, 497–505 (2010)
9. Ifenthaler, D., Hanewald, R.: Digital Knowledge Maps in Education: Technology-Enhanced Support for Teachers and Learners. Springer, New York (2014)
10. Jaynes, E.T.: Probability Theory: The Logic of Science. Cambridge University Press, Cambridge (2003)
11. Koponen, I.T., Nousiainen, M.: Pre-service physics teachers understanding of the relational structure of physics concepts: organising subject contents for purposes of teaching. Int. J. Sci. Math. Educ. **11**, 325–357 (2013)
12. Koponen, I.T., Nousiainen, M.K.: Modelling students knowledge organisation: genealogical conceptual networks. Phys. A **495**, 405–417 (2018)
13. Koponen, I.T., Nousiainen, M.K.: University students associative knowledge of history of science: Matthew effect in action? Eur. J. Sci. Math. Educ. **6**, 69–81 (2018)
14. Koponen, I.T., Pehkonen, M.: Entropy and energy in characterizing the organization of concept maps in learning science. Entropy **12**, 1653–1672 (2010)
15. Lachner, A., Nückles., M.: Bothered by abstractness or engaged by cohesion? Experts' explanations enhance novices' deep learning. J. Exp. Psychol. Appl. **21**, 101–115 (2015)
16. Naudts, J.: Deformed exponentials and logarithms in generalized thermostatistics. Phys. A **316**, 323–334 (2002)
17. Nesbit, J.C., Adesope, O.O.: Learning with concept and knowledge maps: a meta-analysis. Rev. Educ. Res. **76**, 413–448 (2006)
18. Nousiainen, M.: Coherence of pre-service physics teachers views of the relatedness of physics concepts. Sci. Educ. **22**, 505–525 (2013)
19. O'Donnell, A.M., Dansereau, D.F., Hall, R.H.: Knowledge maps as scaffolds for cognitive processing. Educ. Psychol. Rev. **14**, 71–86 (2002)
20. Ronqui, J.R.F., Travieso, G.: Analyzing complex networks through correlations in centrality measurements. J. Stat. Mech: Theory Exp. **5**, P05030 (2015)
21. Servedio, V.D.P., Caldarelli, G., Butta, P.: Vertex intrinsic fitness: how to produce arbitrary scale-free networks. Phys. Rev. E **70**, 056126 (2004)
22. van Boxtel, C., van Driel, J.: "That's in the time of the romans!" knowledge and strategies students use to contextualize historical images and documents. Cogn. Instr. **30**, 113–145 (2012)
23. Tsallis, C.: Possible generalization of Boltzmann-Gibbs statistics. J. Stat. Phys. **52**, 479–487 (1988)
24. Tsallis, C., Mendes, R.S., Plastino, A.R.: The role of constraints within generalized nonextensive statistics. Phys. A **261**, 543–554 (1998)
25. Tsekouras, G.A., Tsallis, C.: Generalized entropy arising from a distribution of q indices. Phys. Rev. E **71**, 046144 (2005)