

Analizando ementas curriculares usando redes complexas

Analyzing course programmes using complex networks

Suzane F. Pinto¹, Ronan S. Ferreira^{*2}

¹Universidade Federal de Ouro Preto, Programa de Graduação em Engenharia de Computação, 35931-008, João Monlevade, MG, Brasil.

²Universidade Federal de Ouro Preto, Departamento de Ciências Exatas e Aplicadas, 35931-008, João Monlevade, MG, Brasil.

Recebido em 18 de março de 2020. Revisado em 18 de março de 2020. Aceito em 23 de maio de 2020.

Analizamos as grades curriculares do ciclo básico das engenharias em nosso instituto usando ferramentas da física estatística de redes complexas. Naturalmente, uma grade curricular estrutura-se em forma de rede (ordenamento temporal e pré-requisitos). Nessa abordagem, cada tópico dentro de uma ementa é associado a um nó que, por sua vez, são conectados por links que representam a dependência de um certo tema para a compreensão de um outro, em uma disciplina diferente. Como uma grade curricular é uma estrutura tempo-dependente, evoluindo semestre após semestre, propomos um modelo simples para assinalar ligações entre pares de nós levando em conta apenas dois ingredientes do processo de ensino-aprendizagem: a recursividade e o acúmulo de conhecimentos. Como já conhecemos nossas grades curriculares, nosso primeiro objetivo é verificar se o modelo proposto é capaz de capturar as particularidades de cada uma delas e identificar as implicações que diferentes sequenciamentos do setor de física possam ter no aprendizado dos estudantes. Nosso modelo pode ser usado como uma ferramenta sistemática auxiliando na construção de uma grade curricular interdisciplinar, articulando entre os saberes das disciplinas iniciais da graduação em ciências exatas.

Palavras-chave: grade curricular, ensino, sistemas complexos.

We analyze the curriculum of the early common-years of engineering in our institute using tools of statistical physics of complex networks. Naturally, a course programme is structured in a networked form (temporal dependency and prerequisites). In this approach, each topic within each programme is associated with a node, which in turn is joined by links representing the dependence of a topic for the understanding of another in a different discipline. As a course programme is a time-dependent structure, we propose a simple model to assign links between nodes, taking into account only two ingredients of the teaching-learning process: recursiveness and accumulation of knowledge. Since we already know the programmes, our objective is to verify if the proposed model is able to capture their particularities and to identify implications of different sequencing on the student learning in the early years of engineering degrees. Our model can be used as a systematic tool assisting the construction of a more interdisciplinary curriculum, articulating between disciplines of the undergraduate early-years in exact sciences.

Keywords: course programme, learning, complex systems.

1. Introdução

O estudo de redes complexas vem passando diversas áreas do conhecimento [1–4]. Desde investigações em dinâmicas sociais [5, 6], passando pela propagação e controle de epidemias [7–9] e até a linguística [10, 11]. Dessa forma, a teoria de redes complexas se tornou um paradigma para pesquisas interdisciplinares que, por sua vez, são um desafio para o ensino, produção e divulgação técnico-científica. Entretanto, nossas grades curriculares dividem-se em conteúdos específicos, por uma motivação didática, levando em conta o processo de ensino e aprendizagem. Outro ponto são os pré-requisitos: exigências sobre competências adquiridas para a continuidade e o acúmulo de conhecimento. Dessa forma, os pré-requisitos

são, ao menos no contexto da interdisciplinaridade, um mecanismo de conexão entre as diversas disciplinas ofertadas em uma grade curricular.

Por outro lado, o conhecimento de uma fração maior dos conteúdos em uma grade curricular por parte dos professores também é um mecanismo fundamental para a consolidação das conexões entre as disciplinas. Isso porque, por parte dos estudantes, nem sempre é fácil perceber a relevância dos tópicos de uma disciplina em outra. Talvez por essa razão indagações como: “*Quando usarei isso?*” ou “*Qual o por quê de eu estar aprendendo isso?*” acompanham o dia-a-dia estudantil.

Vários estudos objetivam uma melhoria na questão de integração dos conteúdos do ciclo básico [12, 13]. Particularmente, nos cursos de graduação em engenharia, os índices de reprovações nas disciplinas dessa etapa vêm se transformando numa espécie de “cultura do insu-

*Endereço de correspondência: ronan.ferreira@ufop.edu.br.

cesso”. Esse pensamento leva os ingressantes (calouros) a considerarem a repetição de disciplinas como Cálculo *Diferencial e Integral e Física* um fato natural: a regra, sendo a exceção aquele aluno regular no curso em questão [12–14]. Consequentemente, as retensões levam a um número elevado de evasões em alguns cursos, implicando em um número grande de vagas ociosas que são difíceis de serem reocupadas quando em períodos mais avançados dos cursos [15].

Em nossa instituição, a Universidade Federal de Ouro Preto - UFOP, foi criado o programa Pró-ativa, o qual visa a melhoria das condições de oferta dos cursos e disciplinas da graduação e o processo de aprendizado. Isso é estimulado com a própria participação de alunos, os quais são selecionados todos os anos pelo programa. Nosso projeto propôs a exposição dos estudantes selecionados à teoria de redes complexas. A partir daí, o grupo foi instigado a mapear as disciplinas do ciclo básico, construindo uma rede de temas conectando as várias disciplinas.

Uma rede é um grafo [3] em que atribuímos um significado físico a um conjunto de nós (vértices) conectados por arestas, obedecendo alguma distribuição estatística. Em nossa abordagem, cada tópico dentro de cada ementa das disciplinas do ciclo básico é associado a um nó, que por sua vez, são conectados aos pares por arestas (*links*). Esses *links* representam, portanto, a dependência de um certo tema de uma disciplina para a compreensão de um outro em uma disciplina diferente. Dessa forma, podemos construir uma *rede de temas* a partir das conexões entre os tópicos das ementas. Nós usamos as ementas das grades curriculares do ciclo básico ofertadas no *Instituto de Ciências Exatas e Aplicadas - ICEA* - da UFOP, disponíveis em (<http://www.icea.ufop.br>).

Nossos objetivos então são (i) a proposição de um modelo para assinalar as dependências entre os diversos temas de uma ementa curricular; (ii) verificar se o modelo proposto é capaz de capturar particularidades de cada grade estudada. A verificação é possível porque conhecemos previamente as grades. Por fim, (iii) identificar a partir do modelo a influência das diferentes sequências dos temas de física tanto para a grade curricular quanto para o processo de ensino-aprendizagem.

2. Metodologia

Podemos utilizar matrizes de adjacências para representar grafos. Para um grafo do tipo $G(N, E)$, em que N é o número de nós e E o número de arestas, podemos escrever uma matriz do tipo $A : N \times N$ nomeando cada nó com um número inteiro $i = 1, 2, 3, \dots, N$. A matriz de adjacências carrega a informação sobre a existência de uma ligação entre quaisquer dois nós em um grafo. Para um grafo do tipo direcionado, os elementos $\{a_{ij}\}$

da matriz A são definidos da seguinte maneira:

$$a_{ij} = \begin{cases} 1, & \text{se uma ligação de } i \\ & \text{para } j \text{ existe;} \\ 0, & \text{se uma ligação de } i \\ & \text{para } j \text{ não existe.} \end{cases}$$

Um grafo direcionado pode, por exemplo, representar uma rede social tal como o *Facebook* ou o *Instagram*: nem sempre seguimos pessoas que nos conhecem, como é o caso dos seguidores de celebridades. Isso pode ser entendido como uma ligação que preserva um sentido. Portanto, podemos representar essa ligação com uma seta indicando um sentido no *link* entre a celebridade e o fã desconhecido. Por outro lado, para o caso de um grafo não direcionado, podemos escrever o conjunto de elementos de uma matriz $B : N \times N$ a partir de suas entradas $\{b_{ij}\}$, em que

$$b_{ij} = \begin{cases} 1, & \text{se uma ligação entre } i \\ & \text{e } j \text{ existe;} \\ 0, & \text{se uma ligação entre } i \\ & \text{e } j \text{ não existe.} \end{cases}$$

A diagonal de uma matriz de adjacências guarda a entrada dos elementos $\{a_{ii}\}$ (ou $\{b_{ii}\}$), informando se um nó está ligado a ele mesmo. Se isso acontece, ele possui uma *autoligação*. Em geral, $a_{ij} \neq a_{ji}$, mas $b_{ij} = b_{ji}$ sempre, logo, B é uma matriz simétrica.

A mais fundamental das quantidades estudadas em redes é o *grau* de um nó, dado pelo número de conexões que ele possui. A distribuição de graus $p(k)$ em um grafo equivale à fração de nós que possuem grau k :

$$p(k) = \frac{N_k}{N}, \quad (1)$$

uma vez que $\sum_k N_k = N$. O grau k_i de um nó i pode ser obtido a partir da matriz de adjacências de um grafo $G(N, E)$. Supondo G não direcionado, então

$$k_i = \sum_{j=1}^N b_{ij}. \quad (2)$$

Caso ele seja direcionado, teremos que contar o número de arestas que entram (seta apontando para o nó i), k_i^{in} , e que saem, k_i^{out} , em um nó i :

$$k_i^{in} = \sum_{j=1}^N a_{ji} \quad (3)$$

e

$$k_i^{out} = \sum_{j=1}^N a_{ij}. \quad (4)$$

O valor médio dos graus $\langle k \rangle$ de um grafo pode ser obtido a partir da distribuição de conectividade $p(k)$.

Por exemplo, para um grafo não direcionado, o valor médio $\langle k \rangle$ pode ser escrito como sendo:

$$\langle k \rangle = \frac{1}{N} \sum_i k_i = \sum_k kp(k). \tag{5}$$

3. Modelo

Com o mapeamento das ementas das disciplinas em uma matriz de adjacências, podemos associar a cada um dos tópicos, um nó e estabelecer um *link* representando a correspondência de um dado tema em uma certa disciplina para a compreensão do conteúdo em uma outra. Dessa forma, obtivemos uma rede do tipo direcionada, em que os *links* possuem suas terminações com setas. Além disso, nossa rede também apresenta uma dependência temporal, porque, evidentemente, disciplinas de semestres posteriores podem depender de conteúdos apresentados em semestres passados. Portanto, nossa rede de temas cresce a cada período e seus links possuem uma direção.

Na prática, para tratarmos o caráter dependente do tempo entre os nós e construirmos nossa rede de temas, propomos o seguinte modelo:

- (i) Aos tópicos em um mesmo período é permitido a conexão de forma distributiva;
- (ii) Os tópicos em um período φ também podem se conectar com o período $\varphi - 1$ (com $\varphi > 2$), dando assim o ingrediente de recursividade, retratando a ideia do acúmulo de conhecimentos.

A Figura 1 ilustra esse mapeamento dos tópicos das ementas em uma rede não direcionada. Cada tópico é denominado por um número¹.

4. Resultados e discussões

As distribuições que encontramos apresentam demasiadas flutuações, o que torna difícil a análise para propósitos estatísticos. Essas flutuações estão, principalmente, na cauda das distribuições em que $x \gg 1$ ou, equivalentemente, para baixas frequências de ocorrência de x .

¹essa correspondência não é mostrada por brevidade.

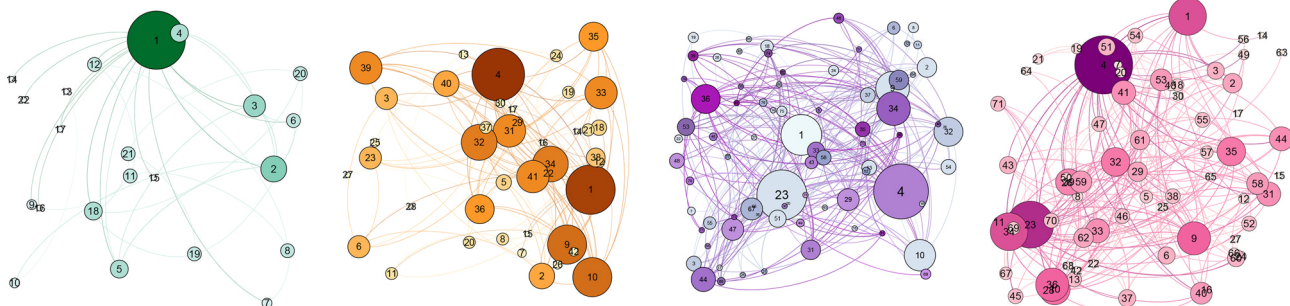


Figura 1: Ilustração do caráter temporal para o aumento da densidade de links e número de nós correspondendo às disciplinas ofertadas em cada período do ciclo básico. Da esquerda para a direita: 1º, 2º, 3º e 4º Períodos do curso de Engenharia Elétrica/UFOP.

Para realizarmos um tratamento estatístico em valores $p(k) \ll 1$, uma estratégia é usar *funções de distribuições acumuladas complementares - CCDF*. De forma geral, uma função CCDF é tal que, no limite contínuo, $f(x) = 1 - \int_{-\infty}^x g(x')dx'$, em que $g(x')$ é uma função densidade de probabilidade, tal que $\int_{-\infty}^{+\infty} g(x')dx' = 1$ e x é uma variável aleatória correspondendo à variável, também aleatória, x' , truncada no intervalo $[-\infty, x]$. Para nossas redes de temas nos basta o caso discreto, em que denotaremos uma função desse tipo por $p_{>}(k)$:

$$p_{>}(k) = \sum_{q \geq k} p(q). \tag{6}$$

No topo da Figura 2 são mostradas as distribuições de entrada (a) $p(k_{in})$ e (b) $p_{>}(k_{in})$, enquanto na base são apresentadas as distribuições de saída (c) $p(k_{out})$ e (d) $p_{>}(k_{out})$ para a rede de temas da Engenharia Elétrica (EE). Em (a), os dados referem-se à distribuição sem o tratamento estatístico CCDF, o qual é mostrado em (b). O melhor ajuste encontrado, seguindo a proporcionalidade $p(k) \propto \exp(-\alpha k)$ é mostrado, tanto em (a) quanto em (b), pelas curvas de traçado contínuo. A mesma sequência de informações é apresentada em (c) e (d) para o caso das distribuições de saída. No decorrer do texto, nos referiremos ao valor do parâmetro α obtido após o tratamento dos conjuntos de dados.

É interessante, neste ponto, fazer uma observação sobre o processo de construção da rede de temas. Nossa primeira investigação foi apenas conectar aqueles temas relacionados, sem nenhuma preocupação com a dependência temporal tampouco pensando no processo de acúmulo de conhecimentos, no qual se baseia, essencialmente, o processo de aprendizagem. Nesse caso, o que encontramos foi uma distribuição exclusivamente aleatória de pontos ordenados ao acaso no plano xy (dados não exibidos). Somente com nossa proposição do modelo apresentado na seção 3 é que obtivemos as distribuições apresentadas na Figura 2.

A Figura 3 apresenta as distribuições de entrada e saída para as três engenharias da UFOP analisadas: Engenharia Elétrica (EE), Engenharia de Computação (EC) e Engenharia de Produção (EP). As curvas de traço conti-

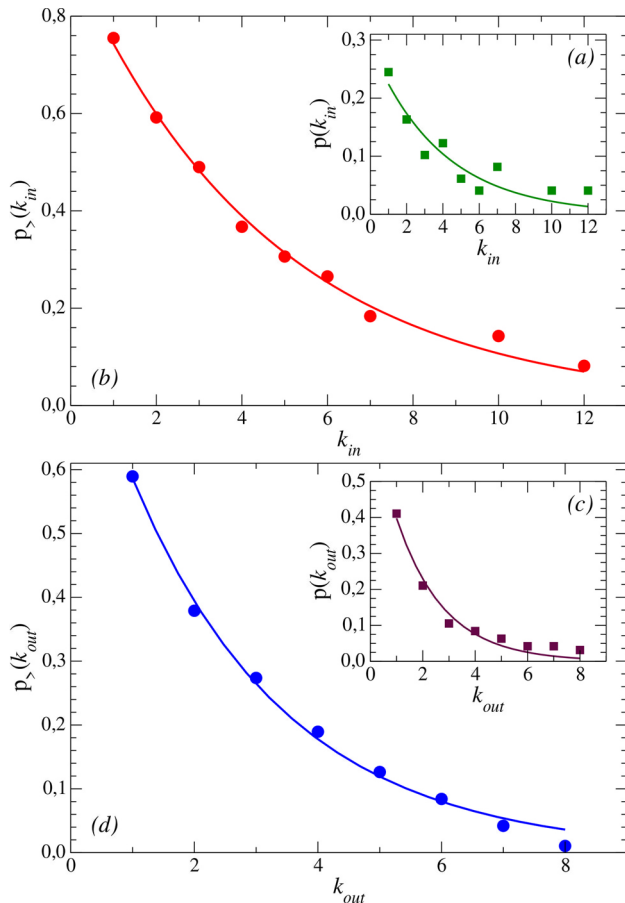


Figura 2: Dos gráficos inseridos (a) e (c) para os respectivos gráficos principais (b) e (d), os dados sofreram um tratamento devido a flutuações para $p(k) \ll 1$. Topo: distribuições de grau para links de entrada. Base: distribuições de grau para links de saída.

nuo são ajustes exponenciais e os valores dos respectivos expoentes são apresentados na Tabela 1.

Para uma comparação, as funções contínuas obtidas nos ajustes exponenciais mostrados na Figura 3 são colocadas em mesma escala na Figura 4. Observe que, apesar de se tratar de dados do ciclo básico das Engenharias e, por isso, esperarmos obter uma superposição das inclinações de todas as Engenharias estudadas, isso só acontece com os dados referentes às Engenharias de Computação e Produção, respectivamente, EC e EP. A superposição não é perfeita devido a ligeiras diferenças nas ementas

Tabela 1: Caracterização quantitativa da rede de temas das engenharias. Correspondência: EE (Engenharia Elétrica), EP (Engenharia de Produção) e EC (Engenharia de Computação).

Característica	EE	EP	EC
Número de nós	96	87	96
Número de links	259	293	300
$\langle k_{in} \rangle$	5,22(1)	4,28(1)	4,14(1)
$\langle k_{out} \rangle$	2,69(1)	3,55(1)	3,47(1)
α_{in}	0,22(1)	0,28(1)	0,29(1)
α_{out}	0,39(1)	0,32(1)	0,33(1)

dos dois cursos. A mais destacada dessas diferenças está no fato da EC ter em seu PPC a disciplina de *Física Moderna (ou Física vol. IV)*, o que não acontece para a EP.

Isso sugere que a inclusão da disciplina de Física IV na grade curricular da EP seria uma proposta ao PPC de curso, a fim de oferecer aos estudantes uma visão atual dos assuntos abordados, implicando uma melhor compreensão dos processos na Engenharia moderna, com destaque para os processos da Física e Engenharia de materiais. Outra interpretação da Figura 3 é que a inclusão da Física Moderna na grade curricular da EP não acarretaria numa sobrecarga aos estudantes do ponto de vista de acúmulo de conhecimento, tampouco do número de pré-requisitos que seriam exigidos para tal disciplina. De outra forma, o desvio entre as inclinações das curvas EC e EP seria evidente, tanto para a distribuição de conexões de entrada quanto para conexões de saída, uma vez que o número de conexões de entrada de um nó está relacionado ao número de pré-requisitos para o tema que este nó representa, enquanto seu número de conexões de saída está associado à relevância deste tema para o andamento do curso.

O desvio mais substancial ocorre entre a Engenharia Elétrica, EE, e as demais. Nesse caso, não acontece a superposição porque também há particularidades em seu projeto pedagógico. De novo, é a setor de física que apresenta as principais diferenças. Por exemplo, no PPC da EE a disciplina de *Eletromagnetismo (ou Física vol. III)* e *Termodinâmica (ou Física vol. II)* estão em um mesmo semestre, enquanto separadas por um período nas demais EC e EP.

Um outro fator é que a disciplina de *Física Moderna* na EE está no semestre imediatamente posterior (φ e $\varphi + 1$) ao período em que é programada a disciplina de *Eletromagnetismo*. Isso porque em sua ementa de *Física Moderna* os primeiros capítulos são dedicados ao estudo de *correntes alternadas*, que por sua vez, possuem relação direta com o estudo dos *circuitos* na disciplina *Eletromagnetismo*. Em nossa abordagem, os nós que representam o tema das *correntes alternadas*, na EE, possuem valores de k_{out} superiores aos mesmos nós na rede da EC, porque nessa engenharia *Eletromagnetismo* e *Física Moderna* estão separadas de um período. O raciocínio inverso se aplica aos valores de k_{in} referentes aos nós que representam o estudo dos *circuitos*. A Tabela 2 esquematiza o posicionamento dessas disciplinas nas grades das três engenharias. Esses fatores levam a uma conectividade média maior para EE mostrando-se uma rede mais rica em pré-requisitos – curva com maior inclinação na Figura 4 (a), embora uma grade desbalanceada se levarmos em conta a relação entre pré-requisitos e relevância de um tema no seguimento de um curso. Isso está relacionado com a razão entre os coeficientes das distribuições de conexões de entrada e saída, em que essa razão r tende a 1 se há uma correlação linear. A saber, $r = 0.6$ para a EE, $r = 0.8$ para a EP e $r = 0.9$ para a EC, esse último

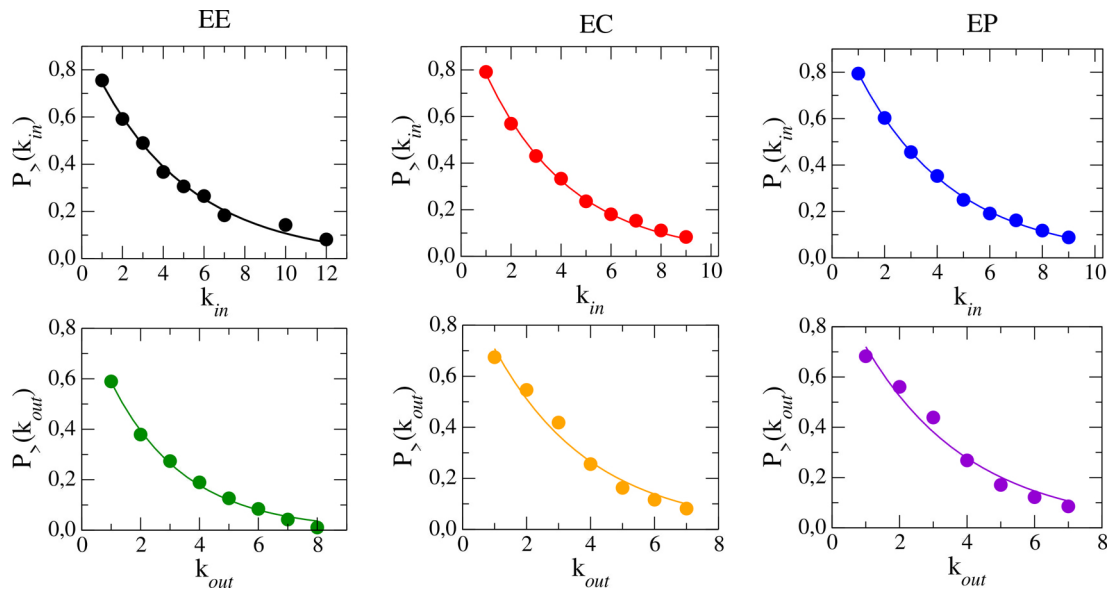


Figura 3: Distribuições $p_{>}(k_{in})$ (topo) e distribuições $p_{>}(k_{out})$ (base) para as conexões entre os tópicos das ementas das três engenharias estudadas. da esquerda para a direita: Engenharia Elétrica (EE), Engenharia de Computação (EC) e Engenharia de Produção (EP). Os valores para os expoentes do ajuste para cada uma das curvas é mostrado na Tabela 1.

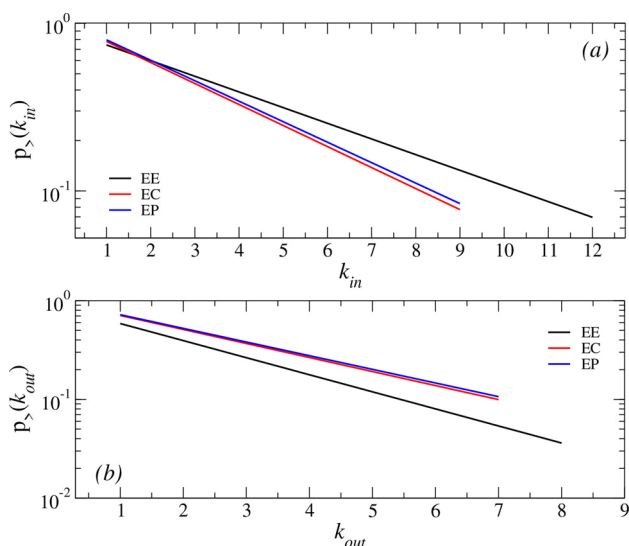


Figura 4: Comparação entre as funções contínuas obtidas na Figura 3. Como se trata da rede de temas para o ciclo básico de três engenharias era esperada uma superposição de curvas. Entretanto, isso não ocorre devido a diferenças no Programa Pedagógico Curricular dos cursos. Nosso modelo conseguiu capturar essas diferenças usando uma abordagem quantitativa.

mostrando-se o curso aqui analisado com a grade mais balanceada do ponto de vista da sequência de apresentação dos temas do ensino de física.

5. Comentários finais

No âmbito do projeto Pró-Ativa/UFOP, que visa pensar a aprendizagem e ensino na graduação, nós propusemos um modelo simples assumindo no processo de aprendi-

Tabela 2: Correspondência: FIS I (Mecânica), FIS II (Eletromagnetismo), FIS III (Termodinâmica) e FIS IV (Física Moderna).

Período	EE	EC	EP
$\varphi = 2$	FIS I	FIS I	FIS I
$\varphi = 3$	FIS II & FIS III	FIS II	FIS II
$\varphi = 4$	FIS IV	FIS III	FIS III
$\varphi = 5$	-	FIS IV	-

dizado os ingredientes de recursividade e acúmulo de conhecimentos. Para testarmos nosso modelo, estudamos as grades curriculares do ciclo básico das três engenharias ofertadas no ICEA/UFOP. Com o mapeamento das correspondências entre os tópicos das grades curriculares em uma rede de temas, pudemos obter distribuições matemáticas descrevendo cada uma dessas grades. Apesar de tratarmos apenas do ciclo básico desses cursos, o PPC de uma dessas engenharias apresenta diferenças comparada às demais. Nosso modelo foi capaz de capturar quantitativamente essas diferenças ao tratarmos as distribuições de conectividade para cada uma das grades curriculares. Nossa análise pode ser estendida tanto para o ciclo profissionalizante quanto para auxiliar no estudo de pré-requisitos em um PPC.

6. Agradecimentos

Os autores gostariam de agradecer o Programa Pró-Ativa/UFOP. RSF gostaria de agradecer o programa Auxílio pesquisador/UFOP, aos professores do DECEA/UFOP, UESPI-Piripiri/Teresina, DFIS/UFPI e Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, no âmbito do processo 424950/2018-9.

Referências

- [1] M.E. Newman, A.L. Barabási e D.J. Watts, *The structure and dynamics of networks* (Princeton University Press, Princeton, 2006).
- [2] A.L. Barabási, *Nat. Phys.* **8**, 14 (2011).
- [3] M.E. Newman, *Networks: an introduction* (Oxford University Press, Oxford, 2018).
- [4] A.L. Barabási, *Sci.* **357**, 138 (2017).
- [5] J. Scott, *Social Network Analysis: a Handbook* (Sage Publications, London, 2017).
- [6] G. Szabó e G. Bunth, *Phys. Rev. E* **97**, 012305 (2018).
- [7] N. Masuda and P. Holme, *Temporal Network Epidemiology* (Springer, Singapura, 2017), p. 179.
- [8] R. Pastor-Satorras e C. Castellano, *J. Stat. Phys.* **173**, 1110 (2018).
- [9] E. Valdano, M.R. Fiorentin, C. Poletto e V. Colizza, *Phys. Rev. Lett.* **120**, 068302 (2018).
- [10] S. Martincic-Ipsic, D. Margan e A. Mestrovic, *Physica A.* **457**, 117 (2016).
- [11] I.G. Torre, B. Luque, L. Lacasa, J. Luque e A. Hernández-Fernández, *Sci. Rep.* **7**, 43862 (2017).
- [12] M.P. Belançon, *Rev. Bras. Ens. Fis.* **39**, 4 (2017).
- [13] C.M.B. Matta, S.M.G. Lebrão e M.G.V. Heleno, *Psicol. Esc. Educ.* **21**, 583 (2017).
- [14] M.F. Barroso e E.B.M. Falcão, em *Anais do IX Encontro Nacional de Pesquisa em Ensino de Física* (Jaboticatubas, 2004).
- [15] A.V.C. Campello e L.N. Lins, em *Anais do XXVIII Encontro Nacional de Engenharia De Produção* (Rio de Janeiro, 2008).