

Modelo de Visualização de Dados para Auxiliar o Professor no Acompanhamento do Desempenho do Aluno e Aprimoramento do *Design* Instrucional com o Apoio de *Learning Analytics*

Douglas Afonso Tenório de Menezes¹, Renato Ely Domingues Silva¹,
Diogo Florêncio de Lima¹, Isabel Dillmann Nunes², Ulrich Schiel¹

¹ Universidade Federal de Campina Grande (UFCG), Universidade Federal do Rio Grande do Norte (UFRN)

douglas2002@gmail.com, {renato.silva, diogo.lima}@ccc.ufcg.edu.br,
bel@imd.ufrn.br, uschiel@gmail.com

Abstract. *The quality of distance learning courses is determined by an appropriate organization of material and learning forms. Therefore it should be taken into consideration, among other factors, the history of success and failure of previous achievements of the course to be offered and the specific profile of the new class of students. This paper proposes an information visualization environment based on the data of the Learning Analytics of a discipline, that supports the teacher to produce improvements in new editions of a course. It should also be possible to monitor runtime student progress and identify potential problems with the Instructional Design of a discipline.*

Resumo. *A qualidade do aprendizado no Ensino a Distância é determinada por uma organização adequada do material e das formas de ensino. Para tal, deverá ser levado em consideração, entre outros fatores, o histórico de sucessos e insucessos de realizações anteriores das disciplinas e o perfil das novas turmas de alunos. Este trabalho propõe a concepção de formas de visualização dos dados obtidos pelo Learning Analytics de uma disciplina que possibilite ao professor a incorporação de melhorias em novas edições de um curso baseado em edições anteriores do mesmo. Também deverá ser possível acompanhar, em tempo de execução, o progresso do aluno e identificar possíveis problemas no Design Instrucional de uma disciplina.*

1. Introdução

O aumento da utilização de Ambientes Virtuais de Aprendizagem (AVA) e outras ferramentas educacionais permitem o avanço da Educação a Distância e também serve de apoio às atividades de ensino presencial [Geller et al. 2004]. Contudo, os dados que são gerados nestes ambientes são pouco explorados para gerar indicadores da qualidade dos programas de formação. Quando esses dados são devidamente tratados e analisados, podem identificar importante conhecimento acerca da adequação dos cursos a seus alunos, identificando fatores que podem influenciar diretamente no processo de aprendizagem [Cavalcanti et al. 2014].

Partindo deste ponto de vista, a análise dos dados computacionais que foram processados e das informações que foram geradas pode ser apoiada por resumos em forma

de visualizações gráficas adequadas, possibilitando uma interpretação e percepção de indicadores importantes dos cursos e seus alunos.

A visualização da informação é uma área que objetiva auxiliar a compreensão de um conjunto de dados, fornecendo um meio de facilitar a análise de dados gerados. Uma técnica de visualização se baseia em uma representação visual, cujo nível de abstração é alto, pois não há relação direta entre os dados visualizados e uma entidade física ou geométrica. O uso da representação visual, interativa e mediada por computador amplia a cognição, ou seja, explora a capacidade de percepção do usuário, permitindo a interpretação e a compreensão das informações apresentadas [Scaife and Rogers 1996].

O objetivo deste artigo é conceber um modelo de visualização de dados que possibilite ao professor acompanhar o progresso do aluno e identificar possíveis problemas no *Design Instrucional* de uma disciplina com os dados obtidos do *Learning Analytics*.

O artigo apresenta conceitos de Learning Analytics na Seção 2. Visualização de dados e trabalhos relacionados sobre visualização de dados educacionais são apresentados na seção 3. Os conceitos sobre Redes de Atividades como ferramenta para o *Design Instrucional* são apresentados na seção 4. Já o modelo de visualização de dados é apresentado na seção 5. A seção 6 apresenta um estudo de caso com a utilização do módulo de visualização de dados proposto neste artigo com seus respectivos resultados. Por fim, as conclusões são discutidas na seção 7.

2. Learning Analytics

Através dos Ambientes Virtuais de Aprendizagem (AVAs), podemos coletar uma grande quantidade de dados de navegação, que nos possibilita fazer diferentes tipos de análises relacionadas ao comportamento do aluno no ambiente, auxiliando na prática pedagógica e até mesmo na melhoria do ambiente. Neste sentido, surgem as Learning Analytics (LA), que é uma área de pesquisa que analisa os dados produzidos por alunos dentro e fora dos AVAs, objetivando melhorar o processo de ensino/aprendizagem.

Learning Analytics são a medida, coleta, análise e relato dos dados de estudantes e seus contextos com o propósito de entender e otimizar o aprendizado e o ambiente em que este ocorre [Siemens 2010]. Este processo pode ser aplicado à disciplinas, cursos e departamentos, fornecendo informações valiosas sobre o que está realmente acontecendo, sugerindo maneiras através das quais os educadores podem melhorar o processo de ensino-aprendizagem, além de indicar quais são os alunos que podem desistir de um curso ou aqueles que precisam de ajuda adicional para melhorar seu desempenho [Siemens and Long 2011].

3. Visualização de Dados Educacionais

A visualização da informação, em sua essência, consiste em representar visualmente um conjunto de dados na forma de uma imagem dentro de um modelo mental, que corresponde a um mecanismo do pensamento pelos quais o homem tenta explicar como são os objetos no mundo real [Ware 2012]. É como se o objetivo fosse projetado para o sistema cognitivo e transformado em um símbolo interno ou numa representação da realidade externa.

Assim como nas representações textuais, as imagens projetadas, na maioria dos

casos, são destinadas a criar uma espécie de correlação visual com o modelo mental do usuário. A visualização da informação também tem vantagens no grau da percepção humana e pode melhorar significativamente tanto a qualidade como a quantidade de informação que é apresentada através das imagens computacionais geradas [Card et al. 1999] [Ware 2012] [Chen 2006].

A cada dia, novas técnicas de visualização estão surgindo para estimular o reconhecimento de padrões, ajudando os usuários na compreensão dos diferentes tipos de níveis cognitivos [Munzner 2014]:

- **Exploração:** quando o usuário não sabe o que está procurando (Descoberta);
- **Análise:** Neste caso, o usuário sabe o que está procurando e a visualização ajuda a determinar se a informação procurada realmente está nos dados (Decisão);
- **Descrição:** É o momento em que os dados são conhecidos, e o usuário deve examinar visualmente para poder descrevê-los (Explicação).

A visualização da informação permite identificar de forma rápida e objetiva se o aluno conseguiu adquirir a habilidade do nível anterior, servindo como pré-requisito para o próximo nível.

O trabalho apresentado utiliza uma combinação de temáticas de três áreas principais: Visualização de Dados, *Learning Analytics*, *Design* Instrucional e Rede de Atividades.

4. Rede de Atividades Básica

A notação de Rede de Atividades (RA), concebida inicialmente por [Torres 1996] é uma notação gráfica que representa atividades de um projeto, suas interdependências e os desenvolvedores envolvidos. É utilizada para modelar e acompanhar o desenvolvimento de um projeto em seus diversos estágios. Este projeto pode ser a construção de um artefato de engenharia ou a aplicação de um curso a distância.

Foram definidos elementos de transição, evento, atividades básica e composta e regras de transição entre os elementos. A cada atividade podem ser associados atributos de recursos, custos e tempos, possibilitando simulações e correções de custo e prazo globais para um projeto, desde que se respeite a estrutura da Rede de Projeto.

Redes de Atividades podem ser aplicadas a qualquer processo dirigido a um objetivo a ser alcançado, ou seja, que visa chegar a um final específico [Nunes and Schiel 2013]. Originalmente a ideia foi utilizar RA para modelar projetos de desenvolvimento de software, mas também podem ser utilizadas em projetos de Arquitetura, Engenharia, Planos Governamentais assim como na aplicação de um curso.

O foco deste documento está no acompanhamento do desempenho do aluno através do suporte oferecido pelo módulo de visualização de dados, proporcionando ao professor a possibilidade de identificação de um problema em uma ou mais atividades de forma mais rápida e objetiva, possibilitando ao mesmo, o melhoramento do *Design* Instrucional (DI) com o apoio do *Learning Analytics*. O conceito de Rede de Atividades mostrou-se adequado para a modelagem de um *Design* Instrucional por ser possível modelar o sequenciamento de atividades, definir tempo das atividades e determinar indivíduos (alunos) que executam a rede.

Uma Rede de Atividades utiliza os elementos gráficos apresentados na Figura 1.



Figura 1. Elementos de uma Rede de Atividades.

5. Modelo de Visualização de Dados Educacionais

O modelo proposto tem como objetivo permitir uma visualização de dados educacionais que possibilite ao professor acompanhar o progresso do aluno e identificar possíveis problemas no *Design* Instrucional de uma disciplina em tempo de execução, assim como no planejamento de novos cursos ou novas edições de cursos já realizados, baseado nos dados obtidos pelo *Learning Analytics* [Rigo et al. 2014].

O modelo em questão é uma aplicação web desenvolvida com a utilização de D3.js, que é uma biblioteca baseada em javascript, que tem uma função muito bem definida: permitir ao desenvolvedor manipular documentos baseados em dados utilizando padrões HTML, SVG e CSS. O módulo de visualização de dados funciona a partir de uma base de dados fornecida por qualquer Ambiente Virtual de Aprendizagem. Para utilizar a base de dados de um determinado AVA, é necessário utilizar uma aplicação auxiliar que está acoplada ao módulo de visualização para poder se comunicar com o banco de dados deste AVA e exportar os dados necessários para um arquivo CSV (Comma-Separated Values).

O primeiro passo para o desenvolvimento do modelo de visualização de dados, foi a identificação de metas de *Design* para a elaboração de relatórios. Em particular, foi necessário identificar as questões a serem respondidas com os dados disponíveis. As perguntas que precisam ser respondidas incluem:

- Qual o Perfil dos Alunos de um curso?
- Como identificar se o DI está adequado?
- Como identificar atividades que impactam de forma negativa?
- Quais os caminhos mais utilizados?
- Quais os tempos mínimos, médios e máximos de realização de uma atividade?
- Como medir a performance do aluno a partir dos pontos ganhos em um curso?

Durante as fases iniciais do planejamento do modelo de visualização de dados, foram identificados alguns desafios que impactaram a forma como responder às perguntas mostradas. Estes desafios incluem:

- Há um grande conjunto de dados que precisa ser apresentado de forma compreensível. O desafio é apresentar estes dados de tal forma que a tela possa ser percebida como um todo, possibilitando ao professor a visualização desses dados como um conjunto completo. Como existe a possibilidade de um DI oferecer vários caminhos a serem percorridos, cada um deles pode ser consultado pelo professor durante a realização do curso;
- Existe um número de elementos no conjunto de dados que são secundários às perguntas apresentadas, mas que fornecem um contexto importante para a consulta. Estes incluem idade, sexo, região, histórico acadêmico anterior e estado civil dos alunos.

A Figura 2 apresenta uma visão parcial do módulo de visualização de dados com base em um DI qualquer. Cada círculo representa uma atividade, e as cores estão divididas de acordo com o desempenho do grupo de alunos que fizeram tal atividade. Foram utilizadas três cores (verde, amarelo e vermelho) para representar o conceito/nota em cada atividade, onde as cores se misturam (através da utilização de gradiente) de acordo com os conceitos adquiridos em cada uma das atividades podendo assumir tonalidades diferentes. Os últimos nós representam os alunos que terminaram o módulo/curso, e suas respectivas cores estão diretamente relacionadas ao desempenho do mesmo ao final do módulo/curso.

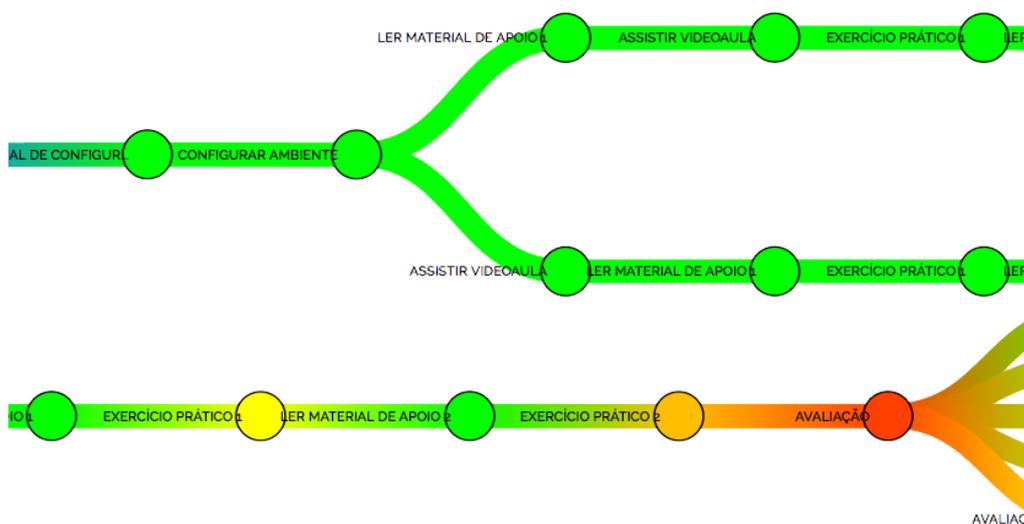


Figura 2. Visualização do andamento do curso.

É possível interagir com a visualização através dos nós. Cada nó representa uma atividade, e ao clicar em um determinado nó, serão apresentadas informações relacionadas a quantidade de alunos que fizeram a atividade; quantos alunos foram bem, quantos foram mal e quantos desistiram (Figura 3).

CAMINHO 1 LER MANUAL DE CONFIGURAÇÃO CONFIGURAR AMBIENTE		
Alunos com nota ≥ 7	Alunos com nota < 7	Desistentes
8	0	0

Figura 3. Visualização do rendimento de uma atividade.

Também é possível visualizar um gráfico de linhas que mostra o desempenho do grupo de alunos até uma atividade em específico. Este tipo de gráfico possui uma grande relevância quando se quer observar uma tendência de notas por grupo de alunos por caminho ou conjunto de atividades.

Além destes, também é possível visualizar, através de um gráfico de Gantt, o desempenho individualizado do aluno com relação a cada atividade, evidenciando como foi feita, quanto tempo levou, o número de tentativas até sua conclusão e o conceito em

cada atividade. É possível identificar os caminhos mais utilizados, os menos utilizados, os tempos mínimo, médio e máximo, apresentando a média de tempo e número de vezes necessário para a realização de uma atividade, com seu respectivo desvio padrão, e ainda fazendo um cruzamento dos dados relacionados aos alunos que foram bem sucedidos e mal sucedidos durante o curso, objetivando comparar o desempenho de cada um desses grupos.

Estes detalhes relacionados ao modelo de visualização de dados serão ilustrados e discutidos na seção de Estudo de Caso.

6. Estudo de Caso

O Serviço Nacional de Aprendizagem Comercial (SENAC) é um dos principais agentes da educação profissional voltado para o setor do comércio de bens, serviços e turismo. Possui uma grande variedade de cursos profissionalizantes relacionados à tecnologia. Dentre eles, o curso de Programação Básica.

O curso de Programação Básica tem como objetivo ajudar o aluno a compreender a lógica de programação antes mesmo de trabalhar com qualquer linguagem de desenvolvimento. O curso é voltado para iniciantes que não dispõem de tempo para assistir aulas presenciais, pois o curso permite aos interessados em programação, desenvolver suas habilidades e raciocínio lógico com autonomia e de acordo com sua agenda pessoal.

O curso é um misto de presencial e à distância, mas em sua maior parte, o curso é feito à distância. O curso possui 4 módulos, e está dividido da seguinte forma:

- Módulo I: Introdução a Algoritmos;
- Módulo II: Dados, comandos e variáveis;
- Módulo III: Estruturas de controle;
- Módulo IV: Estrutura de dados e modularização.

Ao final de cada módulo, é obrigatória a realização de uma avaliação, onde é necessário atingir uma nota maior ou igual a 7 e ter, no mínimo, 70% de frequência nas aulas presenciais.

Com base em entrevistas feitas com instrutores sobre o funcionamento do curso e a elaboração do DI, os mesmos relataram possuir uma grande dificuldade no acompanhamento do progresso do aluno durante a realização do curso, principalmente nos Módulos III e IV. Estes módulos apresentam problemas com relação ao elevado número de reprovação e abandono do curso, fazendo com que a oferta de turmas de programação avançada diminua proporcionalmente ao número de desistentes e reprovados.

Os instrutores reconhecem que existem problemas nestes módulos, mas não conseguem identificar o momento em que os mesmos ocorrem já que o DI possibilita seguir quatro caminhos distintos. Os instrutores já fizeram algumas alterações no DI deste módulo, mas os resultados negativos persistiram.

Para a utilização do módulo de visualização de dados, foi necessário adaptar a base de dados fornecida pelo SENAC para que os dados pudessem ser exportados para o formato e estrutura adequados à sua utilização. Foi criada uma ferramenta para exportar os dados de uma base de dados MySQL para um arquivo CSV (Comma-Separated Values), que é o formato padrão utilizado pelo módulo de visualização de dados. Os dados

foram exportados de 5 entidades diferentes, onde foram selecionados apenas os atributos necessários para o funcionamento do módulo de visualização de dados.

O próximo passo foi adaptar o DI utilizado no curso para o módulo de visualização. Para evidenciar graficamente o DI, foram utilizados os conceitos de Redes de Atividades, e o resultado é apresentado na Figura 4. A representação gráfica é importante para a identificação dos possíveis caminhos existentes no DI.

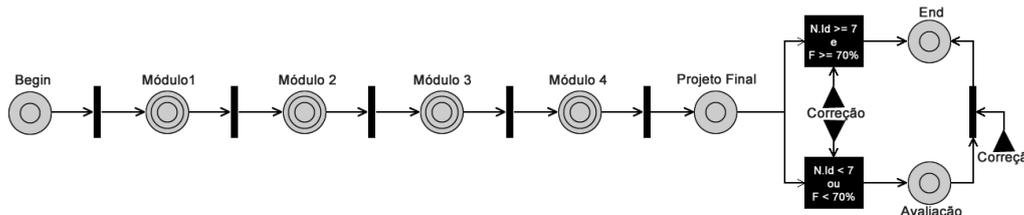


Figura 4. RA do Curso de Programação Básica - SENAC.

A RA da Figura 4 representa o curso como um todo, e está dividido em 4 módulos. De acordo com os relatos apresentados anteriormente, os problemas começam a surgir a partir do Módulo III (Figura 5).

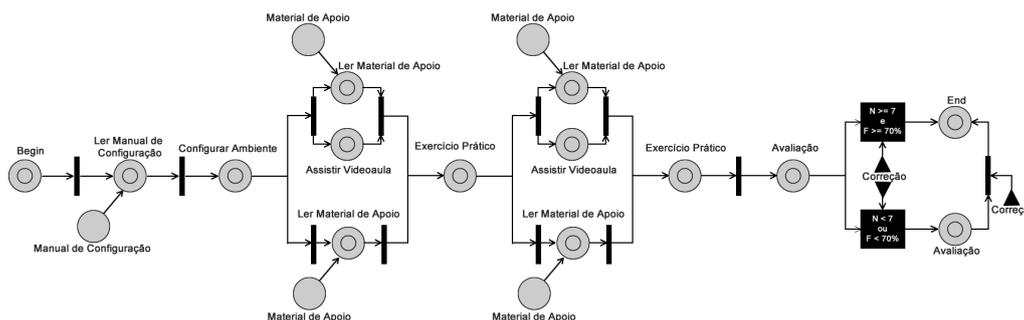


Figura 5. RA do Módulo III do Curso de Programação Básica - SENAC.

O *Design* Instrucional do Módulo III apresentado pela RA da Figura 5, permite quatro possíveis caminhos para a finalização do módulo. Para que o aluno possa concluir o módulo em questão, é necessário que ele obtenha nota igual ou superior a 7 e frequência igual ou superior a 70% nas aulas presenciais. Caso o aluno não consiga, é necessário a realização de mais uma atividade, representada por uma avaliação de recuperação.

Com base na RA do módulo III e com a utilização dos dados fornecidos pelo SENAC, o módulo de visualização de dados evidencia os 4 caminhos possíveis, destacando os dois caminhos que apresentaram problemas com relação ao desempenho dos alunos (Figura 6).

Cada turma de programação básica é iniciada com 20 alunos, logo, é possível observar que mais de 50% dos alunos optaram pela escolha dos caminhos apresentados na Figura 6. Os alunos são representados pelos últimos nós.

O primeiro passo para a validação da ferramenta foi observar se a mesma possibilitava uma rápida identificação das atividades nas quais os alunos possuíam maior dificuldade para sua conclusão. A ferramenta se mostrou muito eficiente neste quesito,

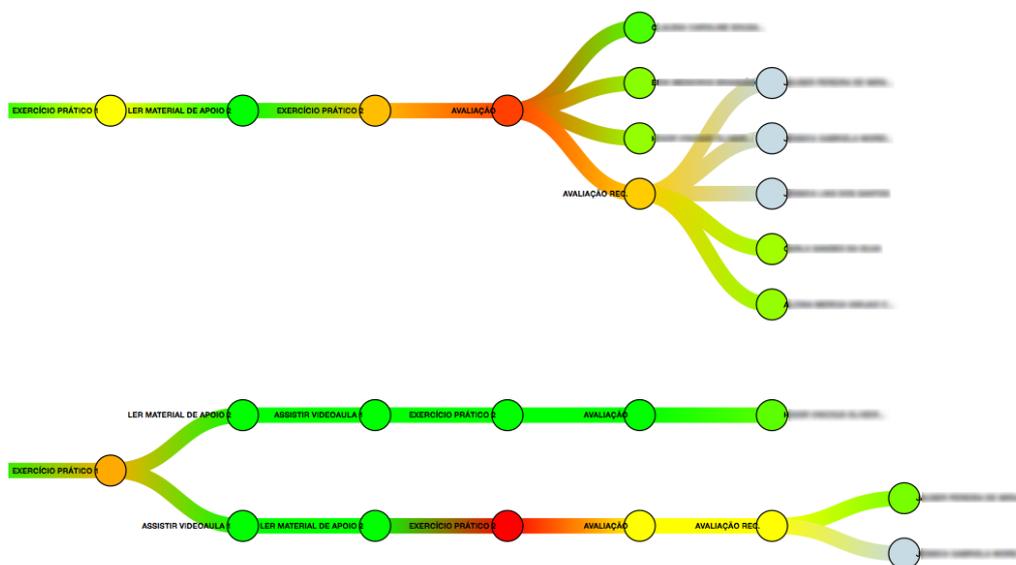


Figura 6. Visualização de dados dos caminhos que apresentam problemas.

pois através da cor vermelha é possível detectar, de forma rápida e objetiva, quais atividades estão apresentando o maior índice de reprovação ou desempenho abaixo do esperado. A cor amarela representa um risco moderado, pois aproximadamente 50% dos alunos que passaram por esta atividade obtiveram notas abaixo de 7.

Além de identificar as atividades que apresentam problemas, o professor tem a possibilidade de obter mais detalhes sobre cada uma delas. Para isto, basta clicar sobre o nó que representa a atividade que o professor deseja analisar e serão apresentados gráficos de desempenho e informações sobre os alunos que passaram pela atividade em questão (Figura 7).



Figura 7. Gráfico de desempenho dos alunos.

As informações relacionadas ao desempenho dos alunos podem ser visualizadas de forma agrupada ou individualmente, onde é necessário selecionar o aluno em questão para obter informações detalhadas sobre cada atividade feita por este aluno. Também é possível identificar o número de alunos que passou por cada uma das atividades, evidenciando os alunos que estão acima da média, os alunos que estão abaixo da média e os alunos que desistiram do curso, através das cores verde, vermelho e cinza, respectivamente.

Informações adicionais podem ser obtidas através dos nós que representam os alunos. Existem duas opções de visualização para o desempenho dos alunos. A primeira é através de um gráfico de barras, que evidencia o desempenho do aluno com relação a cada uma das atividades concluídas.

A segunda opção é através de um gráfico de Gantt que evidencia como as atividades foram feitas pelo aluno, levando em consideração o tempo que levou para concluir, quantas vezes foram acessadas, a ordem em que foram feitas, se foram feitas individualmente ou simultaneamente e suas respectivas notas. Para evidenciar esta funcionalidade, é possível observar na Figura 8, que ao colocar o cursor do mouse sobre a atividade Ler Material de Apoio 1 (Atividade 3) é exibido um quadro com as informações da atividade em questão.

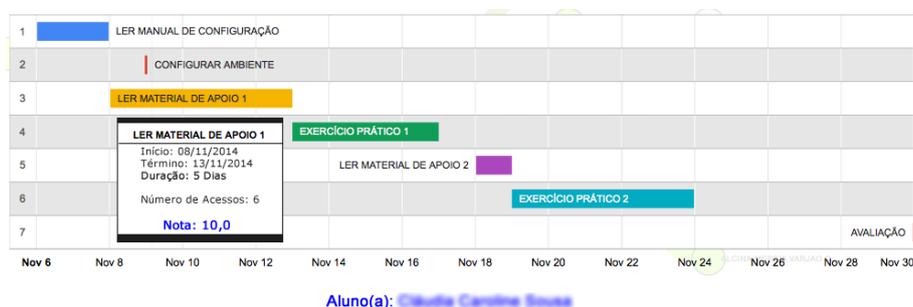


Figura 8. Gráfico de Gantt das atividades realizadas pelo aluno.

7. Conclusão

O objetivo deste trabalho foi o desenvolvimento e criação de um módulo de visualização de dados que possibilitasse ao instrutor acompanhar o progresso do aluno e identificar possíveis problemas no *Design* Instrucional em um curso.

Para ilustrar e testar as visualizações foi criado o *Design* Instrucional de um curso de programação básica no SENAC, que foi representado graficamente pela Rede de Atividades correspondente, que serviu para evidenciar os possíveis caminhos do módulo que apresentava problemas em seu DI.

Após a definição da representação gráfica do DI, foi necessário definir quais atributos seriam utilizados para a utilização do módulo de visualização de dados. Como a ferramenta utiliza um arquivo CSV para o armazenamento de dados tabelados, foi necessário criar uma ferramenta para exportar os dados que se encontravam em um banco de dados MySQL para o arquivo CSV.

À partir destes dados, o módulo de visualização de dados mostra os possíveis caminhos do módulo em questão, evidenciando, através de cores, o desempenho de um grupo de alunos em cada atividade. Além disto, cada nó que representa uma atividade possui informações adicionais sobre um aluno ou grupo de alunos, assim como, os nós que representam os alunos também possuem informações sobre seu desempenho em cada atividade.

Como o principal objetivo na utilização da ferramenta era verificar a possibilidade de identificar, de forma rápida e objetiva, quais caminhos no DI apresentavam mais pro-

blemas e quais as atividades que os alunos possuíam mais dificuldades em concluir com êxito, a ferramenta se mostrou bastante eficiente, pois a identificação do problema é feita praticamente de forma instantânea, evidenciando total satisfação dos instrutores em sua utilização, onde os mesmos se comprometeram a utilizar o módulo de visualização de dados, em tempo de execução, durante a realização de outros cursos.

Como trabalho futuro, o próximo passo será a utilização do módulo de visualização de dados como ferramenta de análise estatística com base nos dados existentes de cursos finalizados de modo que estas informações possam ser cruzadas com os dados de alunos que obtiveram maior e menor índice de aproveitamento, além de algumas características, tais como idade, sexo, região, histórico acadêmico anterior, estado civil, entre outras, possibilitando a elaboração de um DI mais efetivo.

Referências

- Card, S. K., Mackinlay, J. D., and Shneiderman, B. (1999). *Readings in information visualization: using vision to think*. Morgan Kaufmann.
- Cavalcanti, Â. G. G., dos Santos, N. N., Cavalcanti, J. L., and Ramos, A. S. G. (2014). Mineração e visualização de dados educacionais: Identificação de fatores que afetam a motivação de alunos na educação a distância. *ERBASE - XIV Escola Regional de Computação Bahia, Alagoas e Sergipe*.
- Chen, C. (2006). *Information visualization: Beyond the horizon*. Springer Science & Business Media.
- Geller, M., Tarouco, L. M., and Franco, S. R. (2004). Adaptando ambientes virtuais: reunindo educação a distância e estilos cognitivos. In *Anais do Simpósio Brasileiro de Informática na Educação*, volume 1, pages 450–459.
- Munzner, T. (2014). *Visualization Analysis and Design*. A.K. Peters visualization series. A K Peters.
- Nunes, I. D. and Schiel, U. (2013). Utilização de grafo de alcançabilidade para a verificação de design instrucional bem formado. *DesafIE! - II Workshop de Desafios da Computação Aplicada à Educação*.
- Rigo, J. S., Cambrozzi, W., Barbosa, J. L. V., and Cazella, S. C. (2014). Aplicação de mineração de dados educacionais e learning analytics com foco na evasão escolar: Oportunidades e desafios. *Revista Brasileira de Informática na Educação*, 22(01):132.
- Scaife, M. and Rogers, Y. (1996). External cognition: how do graphical representations work? *International journal of human-computer studies*, 45(2):185–213.
- Siemens, G. (2010). 1st international conference on learning analytics and knowledge 2011. Retrieved May, 7:2015.
- Siemens, G. and Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, 46(5):30.
- Torres, J. B. (1996). Uma ferramenta de gerência de projeto–gepro. *Mestrado. UFPB, COPIN, Campina Grande, PB*.
- Ware, C. (2012). *Information visualization: perception for design*. Elsevier.