



Universidade Federal
do Rio de Janeiro

Escola Politécnica

ANÁLISE DOS DADOS COLETADOS DE UMA PLATAFORMA DE EDUCAÇÃO A DISTÂNCIA

Bruno Rafaeli de Miranda Neto

Projeto de Graduação apresentado ao Curso de Engenharia de Computação e Informação da Escola Politécnica da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para a obtenção do grau de Engenheiro de Computação e Informação.

Orientadores: Rosa Maria Meri Leão
Gaspere Giuliano Elias Bruno

Rio de Janeiro
Março de 2018

ANÁLISE DOS DADOS COLETADOS DE UMA PLATAFORMA DE
EDUCAÇÃO A DISTÂNCIA

Bruno Rafaeli de Miranda Neto

PROJETO SUBMETIDO AO CORPO DOCENTE DO CURSO DE
ENGENHARIA DE COMPUTAÇÃO E INFORMAÇÃO DA ESCOLA
POLITÉCNICA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO
GRAU DE ENGENHEIRO DE COMPUTAÇÃO E INFORMAÇÃO.

Examinadores:

Profa. Rosa Maria Meri Leão, Dr.

Prof. Daniel Sadoc Menasche, Ph.D.

Prof. Henrique Luiz Cukierman, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2018

Rafaeli de Miranda Neto, Bruno

Análise dos dados coletados de uma plataforma de Educação a Distância/Bruno Rafaeli de Miranda Neto. – Rio de Janeiro: UFRJ/POLI – COPPE, 2018.

XIV, 34 p.: il.; 29, 7cm.

Orientadores: Rosa Maria Meri Leão

Gaspare Giuliano Elias Bruno

Projeto (graduação) – UFRJ/ Escola Politécnica/ Curso de Engenharia de Computação e Informação, 2018.

Referências Bibliográficas: p. 33 – 34.

I. Maria Meri Leão, Rosa *et al.* II. Universidade Federal do Rio de Janeiro, Escola Politécnica/ Curso de Engenharia de Computação e Informação. III. Título.

Agradecimentos

Agradeço aos meus pais, Helena e Márcio, por todo o esforço e sacrifício para me proporcionarem a melhor educação possível. Por sempre estarem ao meu lado e me apoiarem incondicionalmente, comemorando comigo cada pequena conquista. Vocês são os melhores pais que um filho poderia ter.

Agradeço a minha irmã, Lígia, que sempre esteve ao meu lado e sempre acreditou em mim. Obrigado por todo o carinho.

Agradeço ao meu amor, Thaís, que compartilhou comigo todos os sentimentos dessa trajetória, das angústias e medos às diversas comemorações por cada vitória. Sem ela, o caminho teria sido muito mais difícil. Obrigado pela paciência e por todo o apoio que me deu.

Agradeço a professora Rosa e ao professor Edmundo pela oportunidade de ter feito parte do laboratório LAND e por todo o aprendizado e mentoria nesses anos de graduação.

Agradeço ao meu grande amigo, Gaspare, por todos os ensinamentos, apoio e boas risadas. Também agradeço pela oportunidade de trabalhar em conjunto em seu projeto de doutorado que originou esse trabalho de conclusão de curso.

Agradeço aos meus outros três grandes amigos, Pedro Miranda, Matheus Guedes e Lyang Higa. Juntos vivemos todas as adversidades da graduação e juntos vencemos essa fase.

Resumo do Projeto de Graduação apresentado à Escola Politécnica/COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Engenheiro de Computação e Informação.

ANÁLISE DOS DADOS COLETADOS DE UMA PLATAFORMA DE EDUCAÇÃO A DISTÂNCIA

Bruno Rafaeli de Miranda Neto

Março/2018

Orientadores: Rosa Maria Meri Leão

Gaspere Giuliano Elias Bruno

Curso: Engenharia de Computação e Informação

O serviço de educação a distância vem se tornando uma importante alternativa para estudantes que buscam se profissionalizar. O aumento da procura por esse tipo de serviço traz a necessidade de se estudar a efetividade do aprendizado a distância e o engajamento dos alunos.

O laboratório LAND tem participado de um projeto de educação a distância no estado do Rio de Janeiro. No âmbito desse projeto foi desenvolvido um sistema de videoaulas que atualmente está disponível como um serviço da Rede Nacional de Ensino e Pesquisa (RNP), denominado Videoaula@RNP. O serviço Videoaula@RNP possui mais de 400 videoaulas e é acessado por mais de 40 mil alunos por mês.

Neste trabalho, buscamos entender o comportamento dos estudantes ao assistirem videoaulas de forma a identificar padrões e características que indiquem como construir uma boa videoaula, i.e, uma aula que consiga prender a atenção dos alunos por mais tempo.

Para isso, usamos um conjunto de logs, coletados durante cinco anos (2012 - 2017), com as ações realizadas pelos alunos (pausa, avanço, retrocesso, etc) durante uma videoaula.

A primeira etapa do trabalho foi processar e organizar os dados dos logs em um banco de dados para facilitar a obtenção das métricas de interesse. Quatro métricas principais foram usadas para as análises: (1) SessionTime - duração da sessão do aluno. (2) PlayTime - tempo em "play", que é o tempo que o aluno executa o comando "play". (3) Videotime - tempo em "play"sem repetição. (4) Fração do número de visualizações por fração do tempo de vídeo.

Em uma segunda etapa, analisamos as métricas considerando vários cenários. A partir dos resultados encontrados, propomos uma reflexão acerca da metodologia de ensino utilizada nas videoaulas.

Palavras-Chave: educação a distância, videoaulas, aprendizado, CEDERJ .

Abstract of the Undergraduate Project presented to Poli/COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Computer and Information Engineer.

ANALYSIS OF DATA COLLECTED FROM A DISTANCE EDUCATION PLATFORM

Bruno Rafaeli de Miranda Neto

March/2018

Advisors: Rosa Maria Meri Leão

Gaspere Giuliano Elias Bruno

Course: Computer and Information Engineering

The distance education service has become an important alternative for students who seek to become professionalized. The increase in demand for this type of service brings the need to study the effectiveness of distance learning and the engagement from the students.

The LAND laboratory has participated in a distance education project in the state of Rio de Janeiro. In the scope of this project, a video lecture system was developed, which is available as a service of the Brazilian National Research and Educational Network(RNP), called Videoaula@RNP. The Videoaula@RNP service has more than 400 video lectures and is accessed by more than 40 thousand students per month.

In this work, we seek to understand students' behavior when watching video lectures in order to identify patterns and characteristics that indicate how to construct a good video lecture, that is, a class that can hold students attention for a longer period of time.

For this, we used a set of logs, collected during five years (2012-2017), containing students actions (pause, forward, rewind, etc.) during a video lecture.

In the first part of the work the logs were processed and organized in a database to facilitate the computation of the metrics of interest

Four main metrics were used for the analyzes: (1) SessionTime - duration of the student session. (2) PlayTime - time in "play", which is the time that the student executes the "play" command. (3) Video time - time in "play" without repetition. (4) Fraction of number of views per fraction of video time.

In the second part, we analyze the metrics considering various scenarios. From the results found, we propose a reflection about the learning methodology used in the video lectures.

Keywords: distance learning, video lecture, learning, CEDERJ

Sumário

Lista de Figuras	xii
Lista de Tabelas	xiv
1 Introdução	1
1.1 Educação a distância - Uma breve contextualização	1
1.2 A urgência por novas metodologias de ensino	2
1.3 Exemplos de Sistemas de EaD	4
1.3.1 Edx	4
1.3.2 Coursera	5
1.3.3 Descomplica	5
1.4 Objetivos do trabalho	6
2 Videoaula@RNP	7
2.1 O Serviço	7
2.2 Sistema RIO	8
2.3 Funcionamento do sistema	8
2.3.1 Módulos	8
2.3.2 Interação do aluno	9
2.3.3 RIO Composer	10
2.4 Captura de Logs	10
3 Metodologia	11
3.1 Conjunto de dados	11
3.1.1 Armazenamento	11
3.1.2 Filtragem	12
3.1.3 Agregação	12
3.2 Criação de métricas	13
4 Análise dos dados	14
4.1 Ferramentas	14
4.2 Grupos de estudo	15

4.3	Estatísticas Gerais	17
4.3.1	Amostra	17
4.3.2	Horários de acesso	17
4.3.3	Distribuição ao longo do tempo	17
4.3.4	Diversidade	18
4.4	Agrupamento de videoaulas por número de sessões (popularidade) e duração da sessão	19
4.4.1	SessionTime x PlayTime x VideoTime	19
4.4.2	Fração do número de visualizações por fração do tempo de vídeo	21
4.4.3	Sub-agrupamento por engajamento	22
4.4.4	Sub-agrupamento por período do dia	23
4.5	Agrupamento por dias de prova	25
4.6	Agrupamento por duração da videoaula	26
4.7	Agrupamento por disciplinas recentes	27
4.8	Agrupamento por assiduidade dos cursos	29
5	Conclusão	30
	Referências Bibliográficas	33

Lista de Figuras

1.1	Plataforma Edx	4
1.2	Plataforma Coursera	5
1.3	Plataforma Descomplica	6
2.1	Videoaula@RNP	7
2.2	Exemplo de uma videoaula	8
2.3	Torre de Hanoi	9
4.1	Clusterização das videoaulas	15
4.2	Acessos por período do dia	17
4.3	Distribuição de visualizações ao longo do tempo	18
4.4	SessionTime x PlayTime x VideoTime	19
4.5	SessionTime x PlayTime x VideoTime - Alta Popularidade	20
4.6	SessionTime x PlayTime x VideoTime - Média Popularidade	20
4.7	SessionTime x PlayTime x VideoTime - Baixa Popularidade	20
4.8	Fração do número de visualizações por fração do tempo de vídeo - Aula de alta popularidade	21
4.9	Fração do número de visualizações por fração do tempo de vídeo - Aula de média popularidade	21
4.10	Fração do número de visualizações por fração do tempo de vídeo - Aula de baixa popularidade	22
4.11	Fração do número de visualizações por fração do tempo de vídeo - Aula de alta popularidade	22
4.12	Fração do número de visualizações por fração do tempo de vídeo - Aula de alta popularidade	23
4.13	Fração do número de visualizações por fração do tempo de vídeo - Aula de alta popularidade	23
4.14	Fração do número de visualizações por fração do tempo de vídeo, com filtro por período - Aula de alta popularidade	24
4.15	Fração do número de visualizações por fração do tempo de vídeo, com filtro por período - Aula de média popularidade	24

4.16	Fração do número de visualizações por fração do tempo de vídeo, com filtro por período - Aula de baixa popularidade	25
4.17	Fração do número de visualizações por fração do tempo de vídeo, de uma videoaula baseado em datas de prova	25
4.18	Fração do número de visualizações por fração do tempo de vídeo de uma videoaula	26
4.19	Média de tempo que cada videoaula é assistida	27
4.20	Fração do número de visualizações por fração do tempo de vídeo de uma videoaula mais recente, com duração de 11 minutos e 35 segundos	28
4.21	Fração do número de visualizações por fração do tempo de vídeo de uma videoaula mais recente, com duração de 30 minutos	29
4.22	Assiduidade dos cursos	29

Lista de Tabelas

3.1	sessionsTable	12
3.2	dataTable	12
3.3	sessionStatistic	13
3.4	fileStatistic	13
4.1	Números da amostra	17
4.2	Fração média de visualização por duração da videoaula em minutos	27

Capítulo 1

Introdução

1.1 Educação a distância - Uma breve contextualização

A Educação a Distância é a modalidade educacional na qual a mediação didático-pedagógica nos processos de ensino e aprendizagem ocorre com a utilização de meios e tecnologias de informação e comunicação, com estudantes e professores que estão separados fisicamente no espaço e/ou no tempo [1].

É considerado o marco inicial dessa modalidade de ensino um curso por correspondência oferecido na Gazeta de Boston (Boston, EUA), na edição de 20 de março de 1728. No entanto, apesar de este ter sido o início, a EaD se consolidou e ganhou espaço no Brasil e no mundo no século XX, quando alguns cursos por correspondência se consagraram e, posteriormente, surgiu o ensino através dos rádios e da televisão.[2]

Com o avanço dos meios de comunicação e a ampliação do acesso à internet, entre 1988 e 1991, processou-se a informatização e a reestruturação do sistema de tele-educação, possibilitando uma intensa expansão da EaD no formato virtual, que é o mais explorado hoje.

De acordo com dados do último Censo da Educação Superior do MEC (2016), o número de matrículas em cursos a distância cresceu progressivamente nos últimos dez anos, variando positivamente 297,3% e atingindo o número de 1,5 milhão de alunos em 2016, o que representa 18,6% do total de matrículas da educação superior. Essa expansão é muito significativa, posto que em 2006 os cursos na modalidade EaD representavam apenas 4,2% do total.

O crescimento do ensino a distância pode ser explicado por sua viabilidade financeira, redução de custos e praticidade pedagógica. A flexibilidade de horário e a comodidade para estudar e desenvolver as atividades tornam o professor e o estudante mais autônomos, disponíveis e produtivos[3]. Até mesmo alunos de cursos

presenciais muitas vezes optam por não assistirem as aulas tradicionais e passam a assistir as aulas online [4]. Nesse contexto, a EaD tornou-se um instrumento fundamental de promoção de oportunidades, visto que facilitou o acesso à educação e em muitos casos permitiu o aprimoramento do processo educacional.

É importante ressaltar que falar sobre educação a distância na atualidade significa falar do processo de aprendizagem que utiliza como principais mediadores a internet e os microcomputadores, pois são estes os meios utilizados pela grande maioria das instituições. Além disso, outro ponto interessante é que a ferramenta mais utilizada nessa modalidade de ensino é a tele aula (vídeos) em ambientes virtuais de aprendizagem (AVAs), de acordo com dados do último Censo EaD Brasil (ABED, 2016) [5].

1.2 A urgência por novas metodologias de ensino

O método de ensino tradicional, conhecido como expositivo, é predominante ao redor do mundo há mais de 900 anos [6]. Nesse modelo, o professor é o sujeito ativo no processo de aprendizagem, enquanto os alunos são sujeitos passivos, que assistem simultaneamente ao conteúdo apresentado. [7]

Apesar de ter sido considerada eficaz por muito tempo, essa metodologia vem sendo questionada por diversos estudos nos últimos anos. É difícil determinar se foram as mudanças sociais que afetaram a sua efetividade ou se as novas formas de ensinar permitiram perceber que o método padrão nunca foi verdadeiramente eficiente. De qualquer forma, o que começa ficar claro é que as instituições de ensino precisam se reinventar para que o objetivo de máximo aprendizado seja atingido.

Dois pontos chamam atenção nas pesquisas, pois aparecem em grande parte delas: i) a necessidade de conceder ao aluno um papel ativo no processo educacional; ii) a conciliação da aprendizagem com o uso da tecnologia.

De acordo com o Professor Ronaldo Mota, autor do livro Educando para Inovação, o mundo não é mais fordista, ou seja, não exige que sejam formados técnicos competentes, mas sim, que sejam formadas pessoas capazes de inovar e atuar em conjunto com a tecnologia. Além disso, Ronaldo ressalta que o papel do professor será modificado, pois o processo educacional será coletivo. Ele precisará contar com o auxílio de pessoas da área de tecnologia, design e os próprios alunos serão corresponsáveis pelo ensino, pois as aulas deixarão de ser expositivas para serem colaborativas.

O modelo defendido por Ronaldo ressalta a importância da valorização da aprendizagem independente, porque, de acordo com ele, o modelo atual cria um aluno muito dependente do professor e não o estimula a buscar conhecimento por si só, tornando o conteúdo pouco estimulante. Sendo assim, para o professor, a grande

chave seria o aluno assistir ao conteúdo antes das aulas, para que no momento do encontro com o professor e com os colegas ele esteja preparado para discutir sobre o assunto, bem como apresentar outros pontos de vista e colaborar com a aprendizagem. Ou seja, o processo de aprendizado seria um modelo híbrido, envolvendo encontros presenciais e apresentação de conteúdo à distância.

Ronaldo atribui o sucesso da educação online a essa necessidade de mecanismos alternativos de aprendizado, pois, de acordo com ele, o formato tradicional já não conquista o aluno.

O interessante do modelo proposto pelo professor é que ele atende aos dois principais requisitos apresentados pelos estudos: o aluno como um participante ativo do processo educacional e a utilização da tecnologia.

Uma pesquisa realizada por professores da Universidade de Washington [6] [4], em 2014, traz dados que comprovam a eficiência de modelos como esse. A análise revela que universitários submetidos a aulas no formato de palestras são mais propensos à reprovação e tem notas mais baixas do que alunos em contato com métodos de aprendizado mais ativos e estimulantes.

Diante desse cenário, hoje é sabido que uma mudança precisa acontecer. Basta entrar nas salas de aula para perceber o desinteresse de grande parte dos alunos e o grande desafio que existe para os professores tornarem o conteúdo interessante. Além disso, o mercado de trabalho evoluiu em uma velocidade muito maior do que as instituições de ensino e os recém-formados estão despreparados para se adequarem a ele.

É claro que essa transformação não é simples. Se pensarmos no cenário nacional, o Brasil ainda não está preparado para lidar com esses desafios. Um modelo como esse exige livre acesso a computadores e à internet, o que não é nem de longe a realidade de grande parte dos brasileiros. Entretanto, olhando por outra perspectiva, ainda que o investimento inicial seja alto, a longo prazo um modelo híbrido pode ser mais econômico do que o tradicional, e é claro, pode trazer retornos muito mais positivos do que os resultados alcançados hoje.

Nesse sentido, é preciso uma investigação do que é mais ou menos estimulante para os alunos, para que modelos mais eficientes sejam criados. Ainda que nesse momento não seja possível responder a todos os desafios existentes, é importante repensar o método tradicional e pensar em soluções possíveis, que estimulem alunos e professores a buscarem o conhecimento de forma mais prazerosa e eficiente. O modelo de educação a distância não pode ser puramente baseado em um modelo tradicional e pouco eficiente.

1.3 Exemplos de Sistemas de EaD

Cursos de educação a distância estão se tornando cada vez mais presentes no cotidiano, e ao contrário do que muitos pensam, podem possuir qualidade igual ou até mesmo superior a cursos presenciais. Por não possuírem a limitação de espaço físico e dessa forma serem altamente escaláveis, os cursos de EaD conseguem ter preços altamente competitivos e diversos deles já emitem certificado e diploma de conclusão para o aluno.

1.3.1 Edx

Um exemplo dessa tendência foi a criação da plataforma Edx por duas das maiores e mais renomadas universidades do mundo. A **Universidade de Harvard** e a **Massachusetts Institute of Technology (MIT)** desenvolveram a plataforma Edx, de ensino à distância, para disponibilizarem cursos online, muitas vezes gratuitos ou de baixo custo. O projeto foi muito bem recebido pela comunidade e outras grandes universidades como **Berkeley**, **Columbia University** e **Boston University** aderiram ao projeto. De acordo com os números do site oficial da plataforma, existem cerca de 2042 cursos oferecidos e muitos deles oferecem um certificado de conclusão.[8]

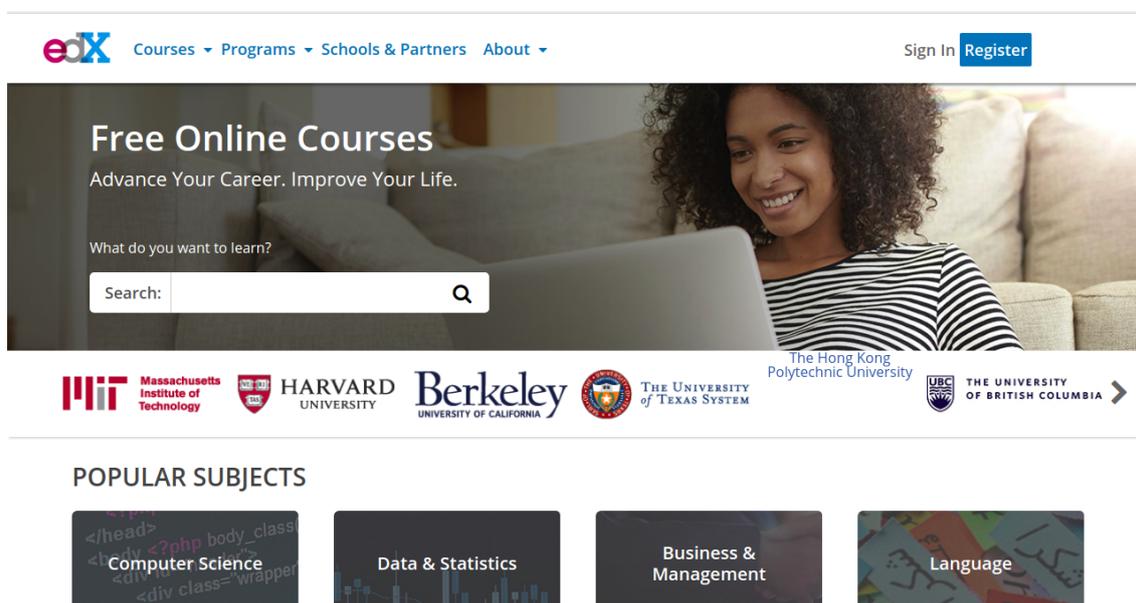


Figura 1.1: Plataforma Edx.

1.3.2 Coursera

Análogo a plataforma Edx, a Coursera foi também criada por renomadas universidades norte americanas como **Stanford**, **Princeton**, **Michigan** e **Pennsylvania**. Obteve o apoio de diversas outras universidades do mundo todo e hoje já possui mais de 2000 cursos à distância, possuindo inclusive cursos de Mestrado com certificação.[9]

No Brasil, a **Fundação Lemann** começou um projeto em 2013, de tradução para o português das videoaulas presentes no Coursera e em 2015 a **USP** e a **Unicamp** passaram a disponibilizar cursos completos na plataforma.[10]

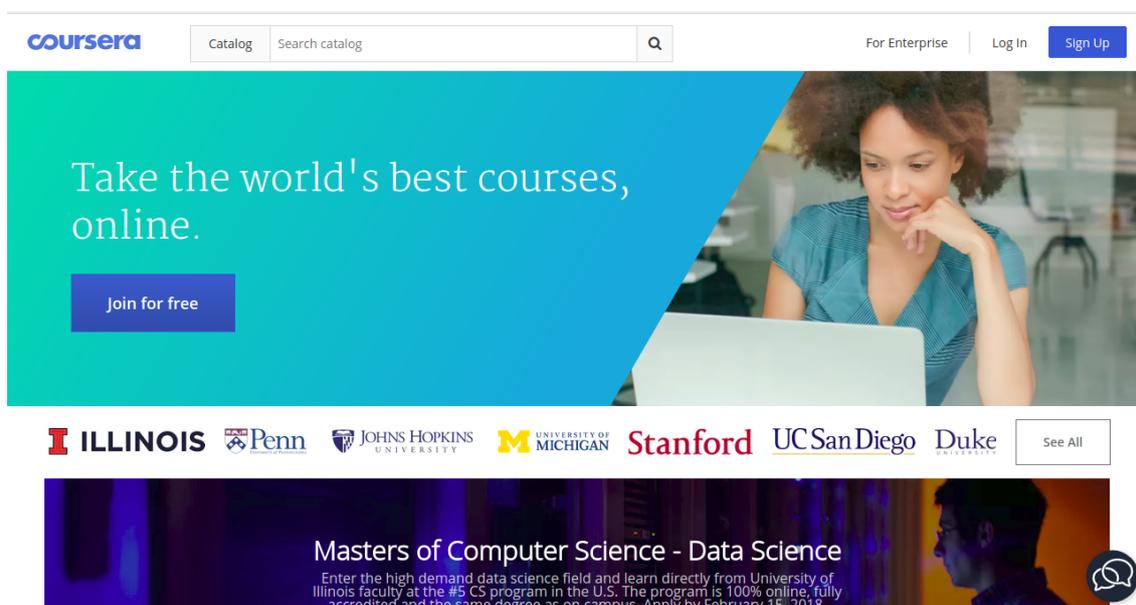


Figura 1.2: Plataforma Coursera.

1.3.3 Descomplica

O Descomplica é uma iniciativa brasileira que surgiu com foco no mercado de pré-vestibular, ENEM e ensino médio. A plataforma é paga e disponibiliza aulas de todas as disciplinas necessárias para o vestibular. O Descomplica já atingiu a casa dos milhões de alunos e hoje é referência no seu setor. [11]



Figura 1.3: Plataforma Descomplica.

1.4 Objetivos do trabalho

A Educação a distância é uma tendência que cresce a passos largos no mundo todo. Com esse crescimento surgem desafios e questionamentos acerca da efetividade e qualidade dessa modalidade. Também crescem as dúvidas quanto a melhor metodologia de ensino, qual a duração ideal de uma aula ou como criar uma aula que funcione bem, na média, para a maioria dos alunos.

O presente estudo surge como uma chance ímpar de analisar os dados coletados de uma plataforma de ensino a distância, que possui milhares de alunos e centenas de videoaulas.

O objetivo aqui proposto é investigar como os usuários do sistema se comportam em diferentes videoaulas, com o intuito de encontrar características que possam ajudar a entender como construir uma boa videoaula, i.e, uma aula que consiga prender a atenção dos alunos e ter alta taxa de assiduidade.

Apartir dos resultados, é proposta uma reflexão sobre a metodologia de ensino utilizada e é dada uma sugestão de nova metodologia.

Capítulo 2

Videoaula@RNP

Neste capítulo será apresentada uma breve contextualização sobre o sistema RIO criado pelo laboratório LAND da COPPE-UFRJ, o serviço Videoaulas@RNP, e o projeto da Fundação CECIERJ/Consórcio CEDERJ de videoaulas do curso de Tecnologia em Sistemas de Computação

2.1 O Serviço

O trabalho proposto nesse projeto utiliza como base o serviço de Videoaula@RNP. Esse projeto se iniciou no laboratório LAND, com a criação do sistema RIO (Random I/O) que possui alta escalabilidade e implementa replicação de informação em múltiplos servidores.

Utilizando esse sistema, a Rede Nacional de Pesquisa (RNP) passou a oferecer um serviço completo e integrado para a criação, armazenamento e distribuição de videoaulas. Todo o conteúdo gravado pode ser acessado a qualquer momento pelo portal Videoaula@RNP : www.videoaula.rnp.br [12]



Figura 2.1: Página oficial do serviço Videoaula@RNP.

2.2 Sistema RIO

O sistema RIO surgiu no ano 2000 no laboratório LAND como um sistema para armazenamento e distribuição de objetos multimídia. No ano seguinte o sistema foi utilizado experimentalmente em um curso semipresencial do consórcio CEDERJ e em 2005 esse mesmo consórcio adotou a solução para o curso de computação, disponibilizando videoaulas em uma rede local para todos os seus alunos. [13]

Nos anos seguintes (2008/2009) os esforços se concentraram em levar o serviço a internet e em 2010 o serviço passou a pode ser acessado via plataforma WEB. Finalmente em 2011, o serviço Videoaula@RNP passou a operar oficialmente.

2.3 Funcionamento do sistema

O sistema de videoaulas permite o compartilhamento de diversos tipos de mídia (áudio, vídeo e texto). A figura 2.2 a seguir mostra um exemplo de uma videoaula em funcionamento, onde pode-se ver os três principais módulos do sistema.

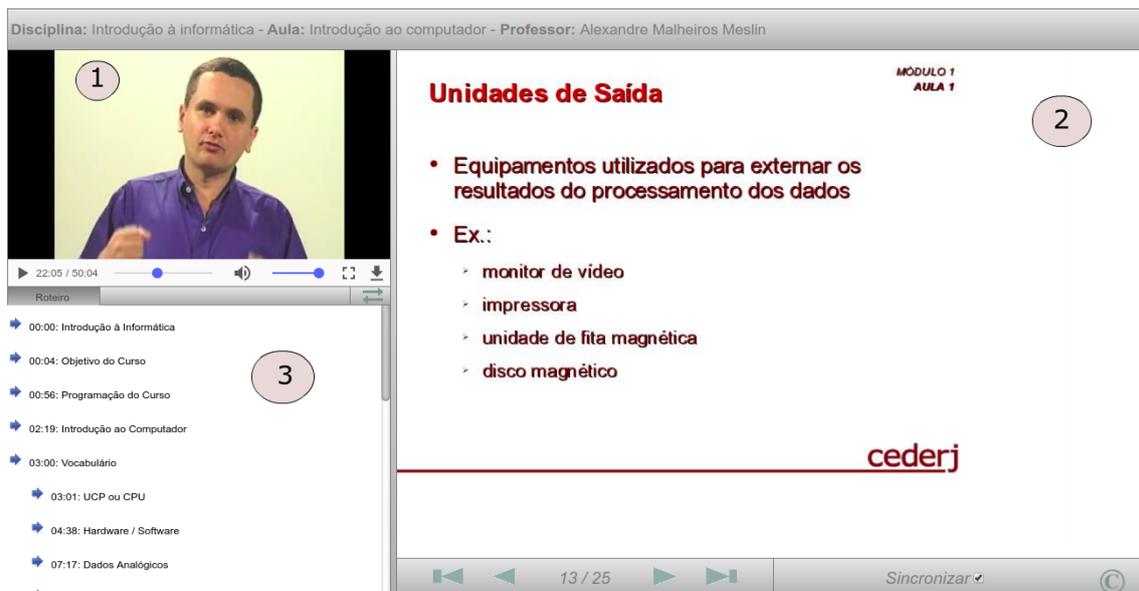


Figura 2.2: Exemplo do funcionamento do sistema em uma videoaula de Introdução à Informática.

2.3.1 Módulos

1. Vídeo

- Módulo onde o aluno pode ter um contato indireto com o professor, como se estivesse assistindo a aula em uma sala normal.

- O aluno pode usar a barra de vídeo para andar para frente ou para trás na videoaula.

2. Slide

- Material de apoio utilizado pelo professor, como um guia para a aula.
- O aluno pode navegar pelos slides, também escolhendo slides para frente ou para trás.

3. Tópicos

- Menu de navegação por assuntos abordados na videoaula.
- Indica o tempo de vídeo e uma legenda para o tópico. Um tópico pode ter sub-tópicos atrelados a ele.
- O aluno pode navegar clicando no tópico que deseja assistir.

Todos os módulos podem ser sincronizados entre si, e ao navegar por qualquer um deles, todos os outros os acompanham.

2.3.2 Interação do aluno

Além de permitir o compartilhamento sincronizado de slides, o sistema permite a exibição de animações que possibilitam que o aluno interaja com a aula. Também é possível compartilhar, por exemplo, um jogo ou exercício interativo.

The screenshot shows a video player interface for a lesson on the Tower of Hanoi problem. At the top, the course information is displayed: "Disciplina: Estrutura de dados - Aula: Recursividade - Professor: Jayme Luiz Szwarcfiter". The main content area features a slide titled "Estrutura de Dados" and "Torre de Hanói". Below the title, it states "• Solução do problema da Torre de Hanói: 3 discos". A diagram illustrates the puzzle with three vertical pegs labeled "Origem", "Destino", and "Trabalho". Three disks are shown on the "Origem" peg. Below the diagram, it says "Total: 7 movimentos" and provides "Mover" and "Voltar" buttons. On the left side, a navigation menu lists various topics with their corresponding video timestamps, such as "17:18: Problema da torre de Hanói", "20:20: Problema com 4 discos", and "27:14: Animação". The video player controls at the bottom show the current time as 27:14 / 45:16 and a "Sincronizar" button.

Figura 2.3: Exemplo de uma animação com o desafio da Torre de Hanoi.

2.3.3 RIO Composer

O RIO Composer é um sistema desenvolvido pelo laboratório LAND que auxilia os professores na construção das videoaulas.

A partir dele, é possível dividir uma videoaula em tópicos e fazer a sincronização dos mesmos com os slides e vídeo.

Também é possível criar animações, formulários e exercícios interativos, em HTML5.

2.4 Captura de Logs

Cada vez que um usuário assiste uma videoaula é gerado um **id único de sessão** que identifica todas as ações realizadas durante a aula. Junto desse id, são registrados informações como IP do usuário, navegador utilizado, sistema operacional, horário de início da sessão e nome da aula e do curso assistido.

O sistema está configurado para capturar informações de **30 em 30 segundos** e também para cada interação realizada pelo usuário durante a videoaula.

As interações capturadas dizem respeito as seguintes ações :

- Iniciar/Pausar o vídeo.
- Navegar pela barra de controle do vídeo.
- Mudar de slide.
- Mudar de tópico usando o menu de navegação por tópicos.

As principais informações guardadas em cada registro são referentes ao **tempo de vídeo decorrido, o estado (play ou pause), o slide e tópico atuais**. Caso o aluno esteja pulando para frente ou pra trás, também são guardados os pontos alvo (target) no qual se deseja chegar.

Ao fim da sessão, é gerado um arquivo de extensão XML contendo todos os registros capturados durante a videoaula.

A lista completa das informações que são capturadas pelo sistema encontra-se no próximo capítulo.

Capítulo 3

Metodologia

Neste capítulo será apresentada a metodologia usada no processamento e filtragem dos dados utilizados no trabalho.

3.1 Conjunto de dados

Como apresentado no capítulo anterior, o sistema registra todos os logs em arquivos de extensão XML.

O trabalho aqui exposto, é baseado nos logs capturados no período de 2012 até 2017, totalizando mais de 40GB de informação.

Esse conjunto de dados contempla informações sobre todas as videoaulas armazenadas no serviço Videoaula@RNP mas esse projeto é relativo apenas as aulas do curso de Tecnologia em Sistemas de Computação do CEDERJ.

3.1.1 Armazenamento

Para uma melhor manipulação dos logs capturados, foi desenvolvido um script em Python, responsável por transportar toda a informação contida nos arquivos XML, para um banco de dados MySQL, que se trata de um banco de código aberto, robusto e que permite a utilização da linguagem SQL para fazer consultas aos dados.

Nesse processo, o banco foi modelado para possuir duas tabelas.

- sessionsTable

Tabela contendo as informações gerais de uma sessão. Possui como chave primária o campo **entrySESSION** que é um identificador único de cada sessão.

- dataTable

Tabela contendo os registros capturados de 30 em 30 segundos ou de cada interação feita pelo usuário. Também possui como chave primária o campo **entrySESSION**.

Campo	Descrição
entryFILE	Nome do arquivo de vídeo.
entryOS	Sistema operacional do usuário
entryVERSION	Versão do Browser
entryBROWSER	Browser utilizado pelo usuário
entryIP	IP do usuário
entrySESSION	Número identificador da sessão
entryTIME	Horário de início da sessão coletado no cliente (máquina do usuário)
entrySERVERTIME_INIT	Horário de início da sessão coletado no servidor de videoaulas (máquina do LAND)
entrySERVERTIME_END	Horário de fim da sessão coletado no servidor de videoaulas (máquina do LAND)
entryCOURSE	Código do curso que a videoaula pertence

Tabela 3.1: sessionsTable

Campo	Descrição
entrySESSION	Identificador da sessão
entryTIME	Horário em que o evento ocorreu
entrySTATE	Estado em que a sessão estava nesse evento (Ex : Play, Pause, Stop)
entryCHANGE	Y ou N (sim ou não). Sim se ocorreu alguma mudança e Não o contrário.
entryVIDEOTIME	Horário em que o vídeo estava.
entryVIDEODOWNLOADEDBYTES	Tamanho do buffer do vídeo.
entrySLIDE	Qual o slide atual que o usuário está vendo.
entryTOPIC	Qual o toico atual que o usuário está vendo.
entryTarget	Para qual ponto do vídeo o usuário está saltando.
entrySERVERTIME	Tempo do servidor.

Tabela 3.2: dataTable

3.1.2 Filtragem

Os logs armazenados no banco podem conter informações inválidas, tanto por problemas durante a captura dos dados, quanto por um comportamento não esperado por parte dos usuários, como por exemplo, a tentativa de acessar um link de uma videoaula não existente.

Para evitar a interferência de dados inválidos, foi criado um script em Python para filtrar apenas os dados de interesse.

Esse processo utiliza algumas regras para separar o que é uma sessão válida ou não.

São consideradas sessões **inválidas**, sessões que apresentam o seguinte comportamento :

- Tempo total de sessão inferior a **1 minuto** ou superior a **180 minutos**.
- Sessões que tenham ficado **1 minuto** ou mais sem nenhum registro.
- Sessões com **30 minutos** ou mais de Pausa.

3.1.3 Agregação

Para facilitar a análise dos dados, foi criado um script para agregar as informações das sessões consideradas válidas. A ideia principal dessa agregação foi a criação de duas novas tabelas que guardam informações pré-processadas dos logs. A partir

dessas novas tabelas foi possível realizar consultas e criar gráficos de maneira mais ágil e otimizada.

As duas novas tabelas são :

- sessionsStatistic

Tabela contendo um resumo de uma sessão. Os campos são id da sessão, dia, mês,ano, período do dia em que a sessão ocorreu, e uma lista contendo o número de visualizações que cada segundo da videoaula teve.

entrySESSION	entryYEAR	entryMONTH	entryWEEKDAY	entryPERIOD	secondVISUALIZATION
20120701114059967	2016	7	6	Morning	[1, 1, 2, 2, 2, 1, 1, 1, 0, 0, 1 ...]

Tabela 3.3: sessionStatistic

- fileStatistic

Tabela contendo um resumo sobre cada disciplina do curso. Os campos são : nome da disciplina, nome do curso no qual a disciplina pertence, número de sessões válidas, número de sessões inválidas e uma lista que informa o número de sessões que tiveram X minutos de duração, sendo X a posição da lista.

entryFILE	entryCOURSE	validSESSIONS	invalidSESSIONS	minuteVISUALIZATION
/cederj/sistemas_comp/ead05020/Aula_001/Aula_001.xml	ead05020	60000	15000	[0, 1000, 150, 300, 500, ...]

Tabela 3.4: fileStatistic

3.2 Criação de métricas

Definiremos a seguir as métricas usadas neste trabalho:

Para cada sessão :

- **SessionTime**

Tempo de duração de uma sessão

- **VideoTime**

Tempo de vídeo assistido sem repetição

- **PlayTime**

Tempo de vídeo assistido com repetição

Para cada videoaula :

- **Fração do número de visualizações por fração do tempo de vídeo**

Considera-se todas as sessões de uma videoaula e contabiliza-se a fração de visualizações que cada segundo teve.

Capítulo 4

Análise dos dados

Neste capítulo serão apresentadas todas as análises feitas a partir dos dados coletados de mais de **500 mil sessões** de videoaulas.

O trabalho aqui apresentado utiliza logs de **463 videoaulas** do curso de **Tecnologia em Sistemas de Computação do CEDERJ**.

4.1 Ferramentas

Todas as análises e gráficos produzidos nesse trabalho foram feitas a partir da linguagem **Python** com as seguintes bibliotecas de apoio.

- **NumPy**

O NumPy é o pacote básico da linguagem Python que permite trabalhar com arranjos, vetores e matrizes de N dimensões, de uma forma comparável e com uma sintaxe semelhante ao software proprietário Matlab, mas com muito mais eficiência, e com toda a expressividade da linguagem. Provê diversas funções e operações sofisticadas.

Fonte : <http://pysciencebrasil.wikidot.com/module:numpy>

- **Matplotlib**

Matplotlib é uma biblioteca Python para criação de gráficos 2D. Ela permite a criação de uma variedade de formatos e também gráficos interativos.

Fonte : <https://matplotlib.org/>

- **MySQLdb**

MySQLdb é uma biblioteca usada para conexão do Python com um banco de dados MySQL. A partir dela é possível realizar consultas SQL dentro do código Python.

Fonte : <http://mysql-python.sourceforge.net/MySQLdb.html>

4.2 Grupos de estudo

Para uma abordagem mais detalhada, mostrou-se interessante separar as videoaulas em diferentes grupos. Ao longo desse capítulo serão apresentadas análises que se utilizam desses grupos para um estudo comparativo.

1. Agrupamentos de videoaulas por número de sessões(popularidade) e duração da sessão

O algoritmo **k-means** foi utilizado para dividir as videoaulas em grupos. As variáveis usadas para fazer o agrupamento foram:

- o número de sessões de uma videoaula, que indicam a sua popularidade.
- a média da duração da sessão, i.e , considera-se todas as sessões de uma videoaula e obtem-se a média da duração.
- o coeficiente de variação da duração da sessão.

Foi utilizado o método **elbow** para escolha do número de **clusters**.

Através do método, verificou-se que as videoaulas podem ser divididas em 3 clusters.

A Figura 4.1 mostra os clusters obtidos após execução do algoritmo k-means.

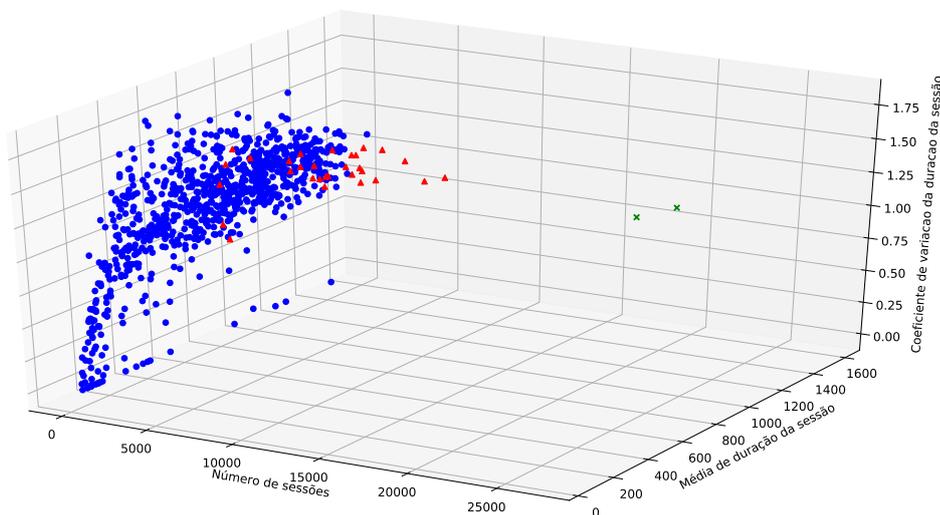


Figura 4.1: Clusterização das videoaulas

2. Engajamento

Nessa modalidade, as videoaulas foram divididas de acordo com a porcentagem de vídeo assistida pelo aluno em relação a duração total da aula.

- 0% a 30% — 30% a 50% — 50% a 70% — 70% a 100%

3. Duração da videoaula

Outra abordagem foi dividir as videoaulas em grupos por **duração da videoaula**. Alguns estudos mostram que a capacidade do aluno de reter informação se esgota em poucos minutos[14]. Surge então o seguinte questionamento : Será que aulas curtas apresentam resultados melhores se comparados a aulas longas?

4. Períodos do dia

Nesse cenário as sessões foram divididas de acordo com o horário em que foram assistidas. O dia foi dividido em 4 períodos.

- **Manhã** - 6h às 12h
- **Tarde** - 12h às 18h
- **Noite** - 18h às 00h
- **Madrugada** - 00h às 6h

5. Datas de provas

Essa abordagem leva em conta as sessões que aconteceram perto de um período de provas de acordo com o calendário acadêmico do CEDERJ. A ideia é investigar se os alunos apresentam comportamento diferente ao estudarem na véspera da prova.

6. Aulas atualizadas

O curso de Sistemas em Computação possui muitas aulas antigas e carentes de atualização. O intuito dessa análise é comparar o comportamento dos alunos ao assistirem videoaulas mais recentes e entender se isso é um fator influenciador.

4.3 Estatísticas Gerais

4.3.1 Amostra

A tabela a seguir apresenta um breve resumo dos dados utilizados nas análises a serem apresentadas nesse capítulo. Cabe ressaltar que por questões de processamento, não foram utilizados todos os registros existentes (que chegam a mais de 2 milhões de sessões). O algoritmo de filtragem dos dados foi bastante rigoroso e foram descartadas todas as sessões consideradas inválidas. No entanto, a amostra aqui utilizada é extremamente significativa e representa bem o comportamento dos dados.

Cursos	26
Videoaulas	463
Sessões	550422

Tabela 4.1: Números da amostra

4.3.2 Horários de acesso

A figura 4.2 mostra a porcentagem do número de sessões por período do dia (**manhã, tarde, noite e madrugada**)

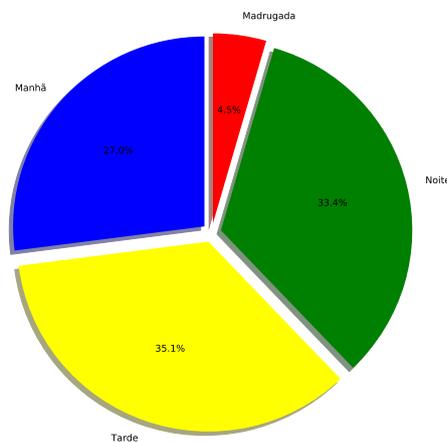


Figura 4.2: Acessos por período do dia

4.3.3 Distribuição ao longo do tempo

Os gráficos da figura 4.3 mostram o número de sessões existentes em cada mês do ano (gráfico inferior) e por dias da semana (gráfico superior). O eixo x representa o período e o eixo y o número de sessões.

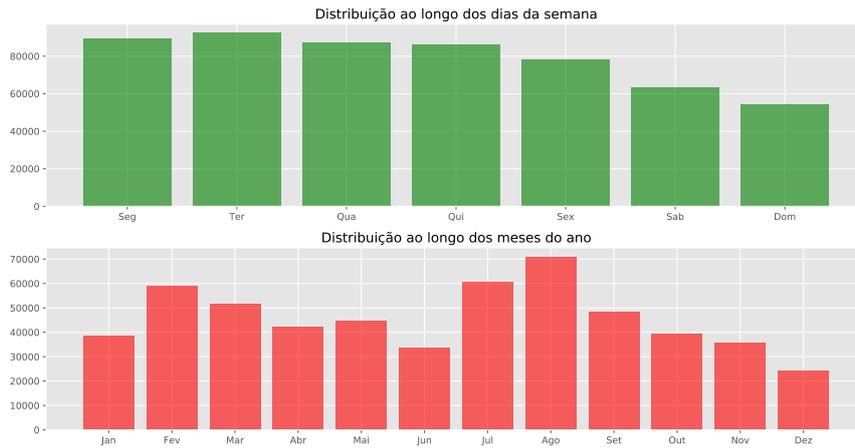


Figura 4.3: Distribuição de visualizações ao longo do tempo.

4.3.4 Diversidade

Apesar de ser um curso para alunos do **Rio de Janeiro**, o sistema videoaulas@rnp é aberto para o público e recebe muitos acessos dos mais diversos lugares do mundo.

A partir de um script em Python que converte o endereço **IP da sessão** em **localização geográfica**, chegou-se a um número próximo de **50%** dos acessos originados de fora do Rio de Janeiro. Esse número mostra como o serviço videoaulas@rnp é abrangente e possui os mais diversificados usuários.

4.4 Agrupamento de videoaulas por número de sessões (popularidade) e duração da sessão

4.4.1 SessionTime x PlayTime x VideoTime

O gráfico 4.4 apresenta as métricas **SessionTime**, **PlayTime** e **VideoTime** no formato de gráfico em barras. O eixo x representa uma **duração em minutos** e o eixo y o **número de sessões** que possuem determinada duração. Esse gráfico foi construído a partir de todas as sessões existentes no banco.

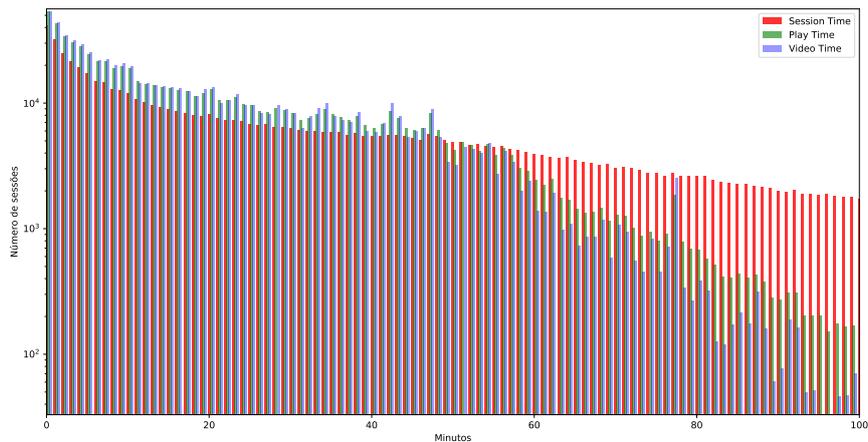


Figura 4.4: SessionTime x PlayTime x Videotime - Todas as sessões

Um comparativo entre essas três métricas mostra que na média, os usuários que assistem as aulas por até **60 minutos**, o fazem sem repetirem partes, i.e, a maior parte das sessões apresentam um VideoTime similar ao PlayTime. A medida que o tempo de vídeo assistido aumenta o SessionTime de muitas sessões também aumenta. No entanto, o número de sessões com PlayTime e VideoTime maiores que 60 minutos, diminui de duas ordens de grandeza. Isso indica que sessões com duração maior que 60 minutos apresentam maior número de repetições e/ou maior tempo em pausa.

Os gráficos das figuras 4.5, 4.6 e 4.7 apresentam as mesmas métricas da figura 4.4 mas separados de acordo com a popularidade, sendo o 4.5 de **alta popularidade**, o 4.6 de **média popularidade** e o 4.7 de **baixa popularidade**.

Se comparados com o gráfico 4.4 pode-se ver que o comportamento é semelhante, para as aulas de média e baixa popularidade.

O gráfico da figura 4.7 onde são apresentados os resultados das videoaulas de alta popularidade possui um comportamento diferente dos demais. O número de sessões com VideoTime e PlayTime inferiores ao SessionTime começa a decrescer após 80 minutos. Essa gráfico utiliza dados de apenas duas videoaulas que possuem no máximo 80 minutos de duração. O resultado mostra que as aulas consideradas

populares conseguiram prender a atenção dos alunos por mais tempo que as de média e baixa popularidade.



Figura 4.5: SessionTime x PlayTime x Videotime - Alta Popularidade

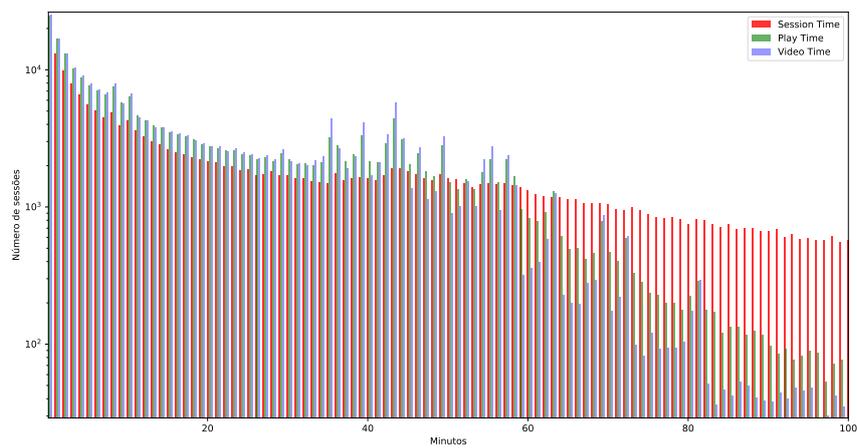


Figura 4.6: SessionTime x PlayTime x Videotime - Média Popularidade

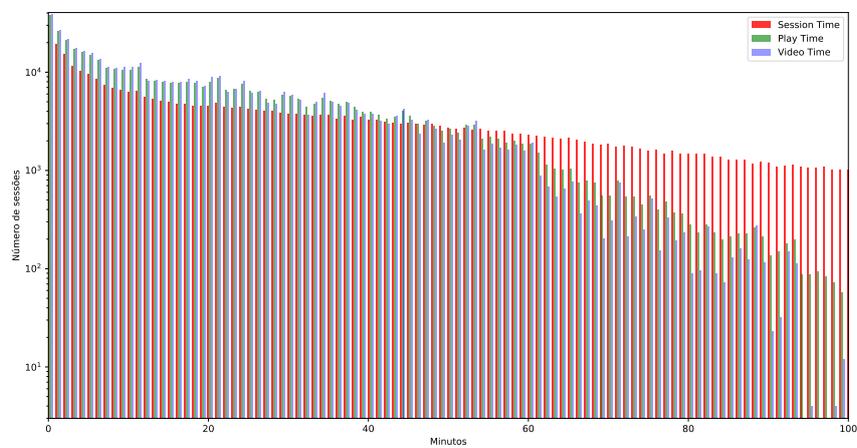


Figura 4.7: SessionTime x PlayTime x Videotime - Baixa Popularidade

4.4.2 Fração do número de visualizações por fração do tempo de vídeo

Os gráficos das figuras 4.8, 4.9 e 4.10 a seguir, possuem no eixo x cada **segundo de vídeo** e no eixo y o **número de visualizações** que cada segundo teve. Ambos os eixos estão normalizados. As linhas verticais em vermelho marcam o começo dos **tópicos** de uma aula.

O intuito dessa análise é mostrar o comportamento dos usuários ao assistirem uma videoaula. Para isso, foram selecionadas três aulas com diferentes popularidades (**alta popularidade, média popularidade e baixa popularidade**).

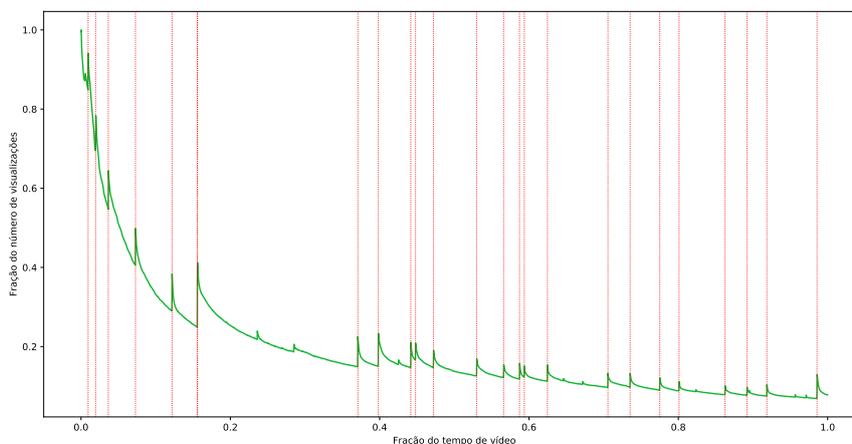


Figura 4.8: Fração do número de visualizações por fração do tempo de vídeo - Aula de alta popularidade

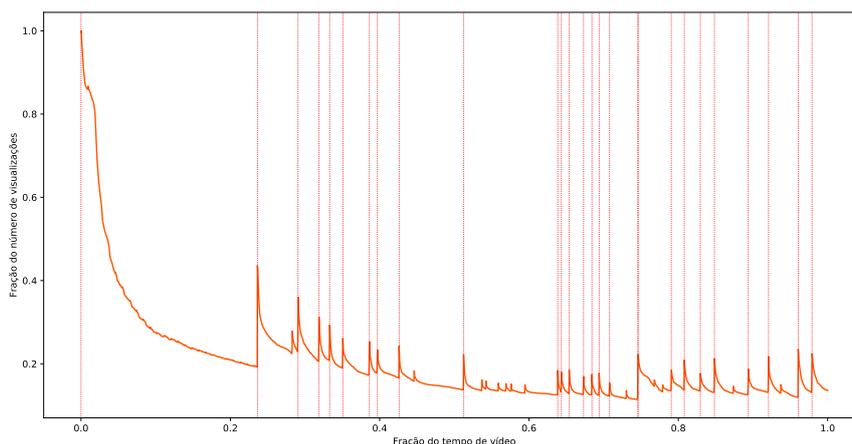


Figura 4.9: Fração do número de visualizações por fração do tempo de vídeo - Aula de média popularidade

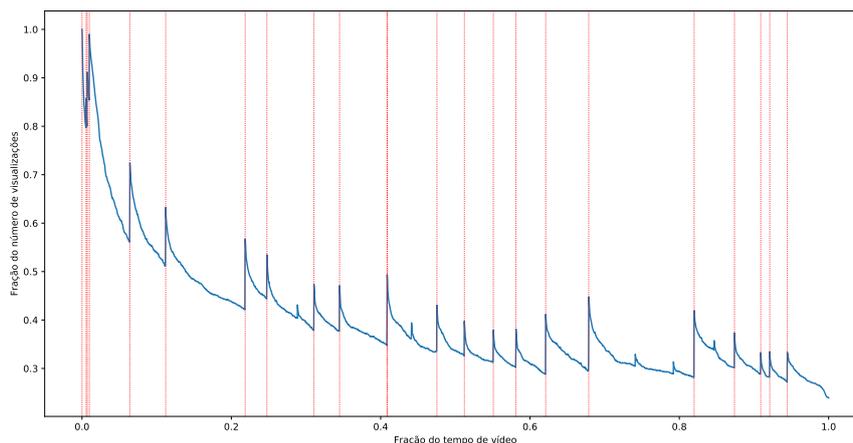


Figura 4.10: Fração do número de visualizações por fração do tempo de vídeo - Aula de baixa popularidade

Os gráficos mostram que o comportamento padrão dos usuários é de navegar pulando entre os tópicos. Cada início de tópico recebe uma grande quantidade de visualizações que decai muito até o final desse mesmo tópico. Além disso, o número de visualizações dos segundos iniciais de uma videoaula é muito superior ao dos segundos finais, o que mostra a alta taxa de evasão durante a aula.

4.4.3 Sub-agrupamento por engajamento

Os gráficos das figuras 4.11, 4.12, 4.13 a seguir mostram uma visualização parecida com a apresentada anteriormente mas dessa vez, separada em 4 curvas diferentes, cada uma representando a porcentagem de vídeo que o aluno assistiu. Cada gráfico é relativo a um grau de popularidade (**alta popularidade, média popularidade e baixa popularidade**).

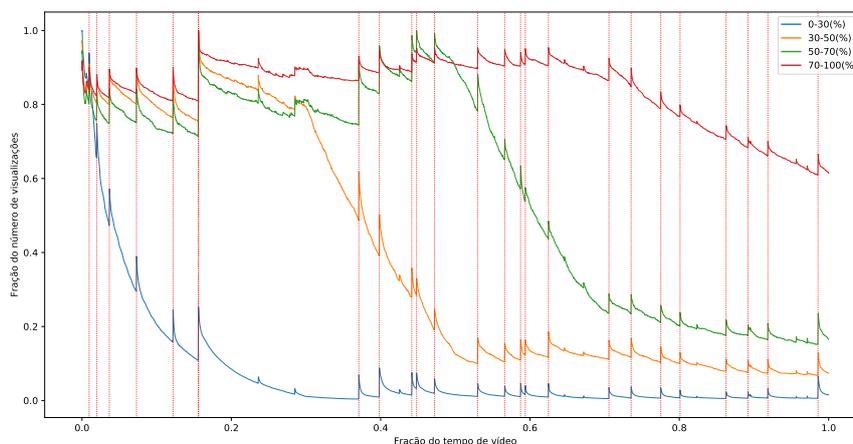


Figura 4.11: Fração do número de visualizações por fração do tempo de vídeo - Aula de alta popularidade

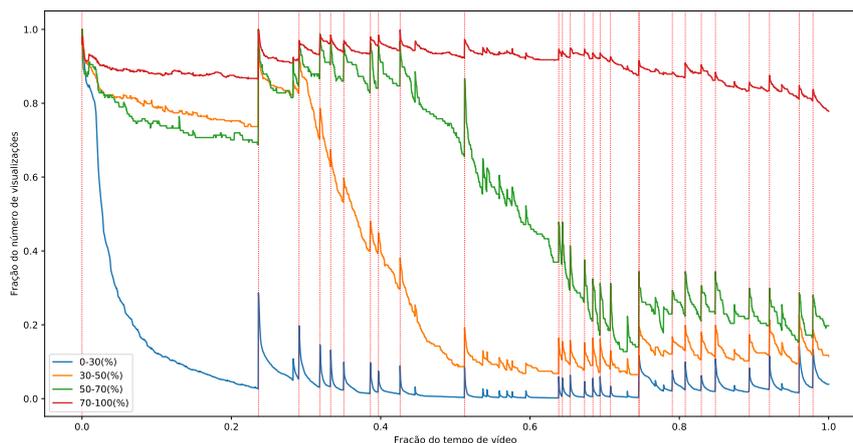


Figura 4.12: Fração do número de visualizações por fração do tempo de vídeo - Aula de média popularidade

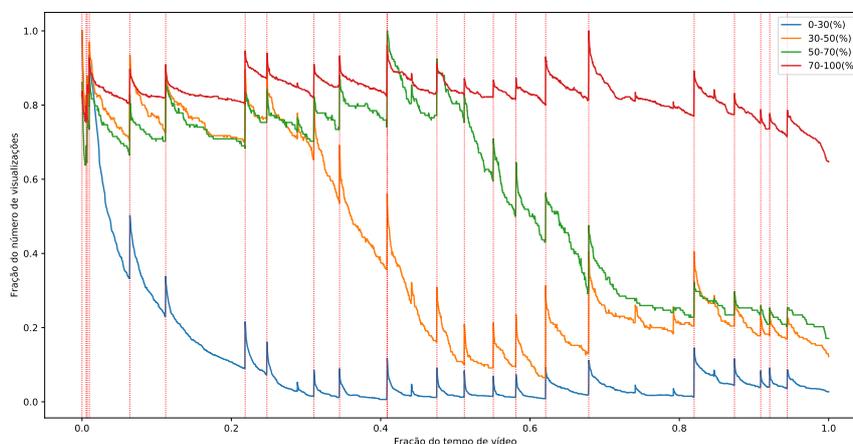


Figura 4.13: Fração do número de visualizações por fração do tempo de vídeo - Aula de baixa popularidade

Esse agrupamento foi usado com o intuito de comparar o comportamento de alunos que se engajaram mais ou menos ao assistirem a videoaula. Considera-se como medida de engajamento o percentual de videoaula assistida pelo aluno. Pode-se observar que o número de visualizações dos alunos que assistiram mais de 70% da videoaula, decresce em torno de 20% do início até o final da aula. No entanto, o número de visualizações dos alunos que assistiram menos de 70% da videoaula, decresce em até 90% até o final da aula.

A partir dos gráficos podemos observar também o mesmo comportamento já observado anteriormente, um maior interesse no início do tópico.

4.4.4 Sub-agrupamento por período do dia

Continuando a análise da videoaula anterior, foi aplicado um filtro por período do dia em que a aula foi assistida. O intuito desse estudo é entender se o horário em

que o usuário assistiu a videoaula poderia interferir no padrão de estudo.

As figuras 4.14, 4.15 e 4.16 a seguir, apresentam a fração do número de visualizações de cada segundo de vídeo pela fração de tempo do vídeo. Análogo aos gráficos anteriores essa visualização mostra que o comportamento dos usuários durante a videoaula não difere de acordo com o período do dia em que foi assistido e pode-se observar novamente o comportamento de navegação por tópicos. Além disso, a popularidade da aula também não se apresenta como um fator de peso.

Vale ressaltar aqui que o curso de Tecnologia em Sistema de Computação é muitas vezes acessado por alunos que trabalham e estudam e por isso a investigação se isso poderia afetar o comportamento dos mesmos.

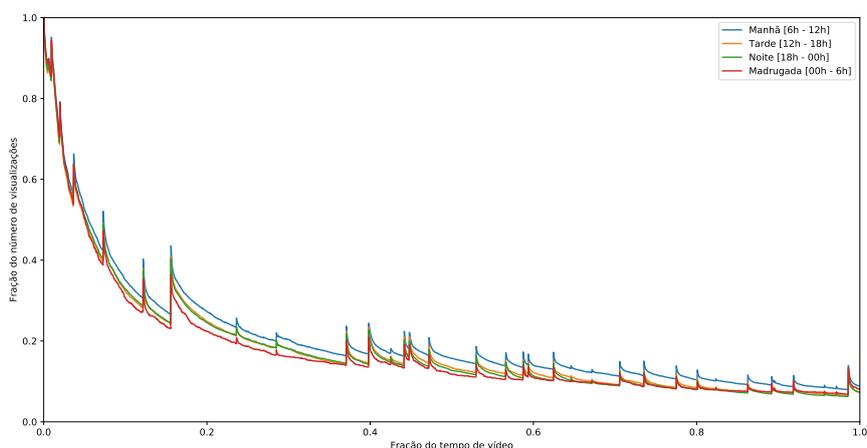


Figura 4.14: Fração do número de visualizações por fração do tempo de vídeo, com filtro por período - Aula de alta popularidade

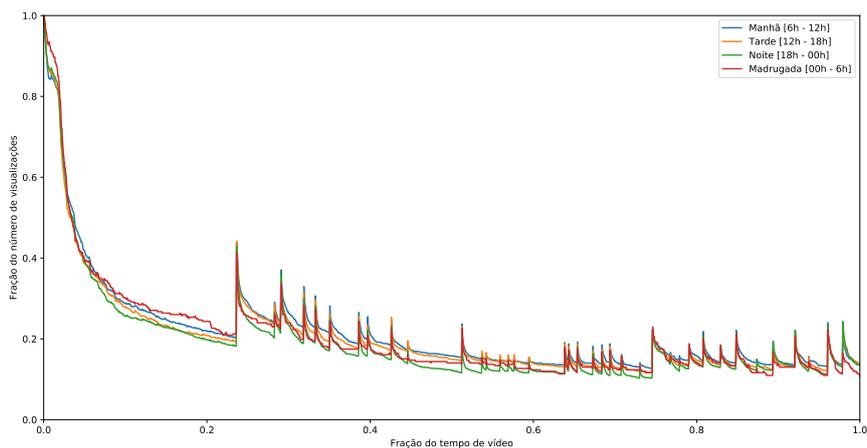


Figura 4.15: Fração do número de visualizações por fração do tempo de vídeo, com filtro por período - Aula de média popularidade

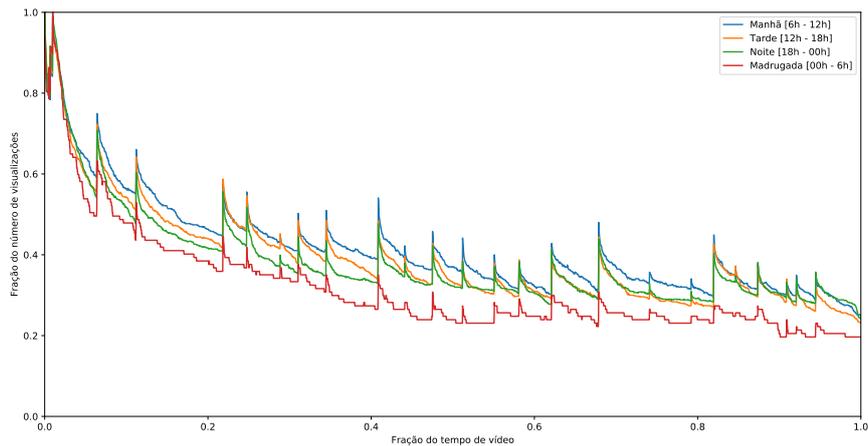


Figura 4.16: Fração do número de visualizações por fração do tempo de vídeo, com filtro por período - Aula de baixa popularidade

4.5 Agrupamento por dias de prova

Apesar de ser um curso de educação a distância, os alunos precisam fazer provas presenciais para serem aprovados nas disciplinas. Baseado no calendário com as datas das provas, o seguinte gráfico foi produzido, com o intuito de comparar o comportamento de estudo das sessões ocorridas até 3 dias antes do dia da prova e das sessões dos dias normais.

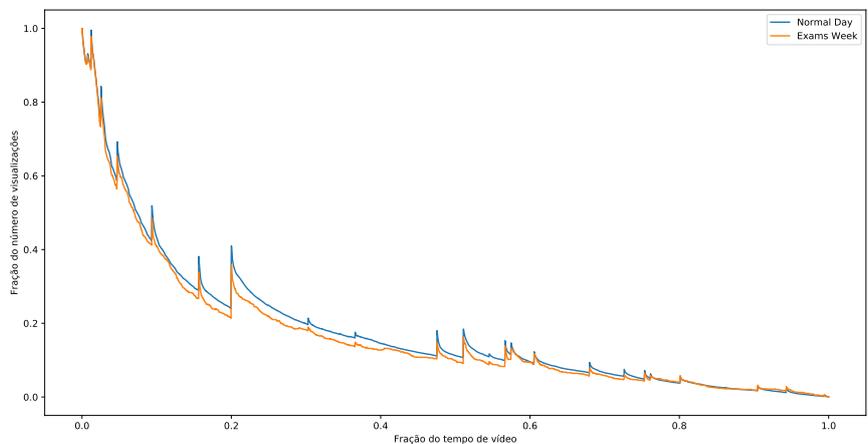


Figura 4.17: Fração do número de visualizações por fração do tempo de vídeo de uma videoaula, baseado em datas de prova

O gráfico da figura 4.17 mostra que as curvas são praticamente idênticas e não possuem diferenças de comportamento.

Esse estudo não pode ser considerado 100% eficaz visto que o número de visualizações próximos da data de prova foram consideravelmente menores que o dos dias restantes.

4.6 Agrupamento por duração da videoaula

Muito se tem discutido sobre o tempo de duração ideal que uma aula deve ter. Alguns estudiosos do tema defendem que os estudantes perdem a atenção após poucos minutos decorridos de aula, e dessa forma passam a ter baixa produtividade no restante da mesma. O especialista em neuroeducação, Francisco Moura defende por exemplo, que é mais proveitoso assistir 50 aulas de 10 minutos do que 10 aulas de 50 minutos cada. [14]

Com base nessa teoria, o intuito dessa análise é entender se aulas mais curtas apresentam comportamento diferente de aulas mais longas.

No gráfico da figura 4.18 pode-se ver a fração do número de visualizações em 3 aulas distintas, com **10**, **30** e **60** minutos de duração. O comportamento de navegação por tópicos se mantém, mas a queda de visualizações de um tópico para outro é mais lenta na aula mais curta em comparação com a a aula de maior duração. Esse comportamento pode indicar que os alunos tendem a se manter na aula por saberem que a duração é menor.

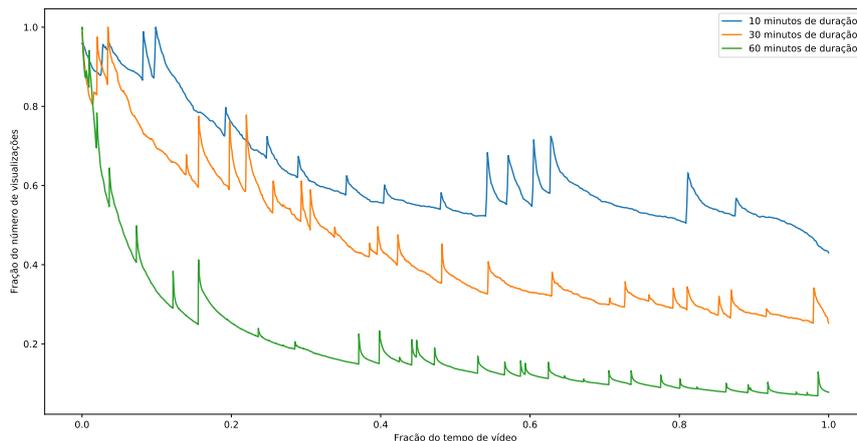


Figura 4.18: Fração do número de visualizações por fração do tempo de vídeo de uma videoaula

A próxima análise gráfica mostra qual a média de tempo em minutos sem repetição (VideoTime) que cada videoaula é assistida (**eixo y**) de acordo com o tempo total que ela possui (**eixo x**).

Essa visualização mostra que na média os usuários do sistema não assistem mais do que **30 minutos** de aula, independente do tamanho da aula.

O resultado é muito interessante, pois, 68% das videoaulas do curso de Sistemas de Computação possuem mais do que 30 minutos de duração, e dessa forma, muitas delas podem não estar atingindo todo o seu potencial e sendo bem aproveitadas pelos alunos.

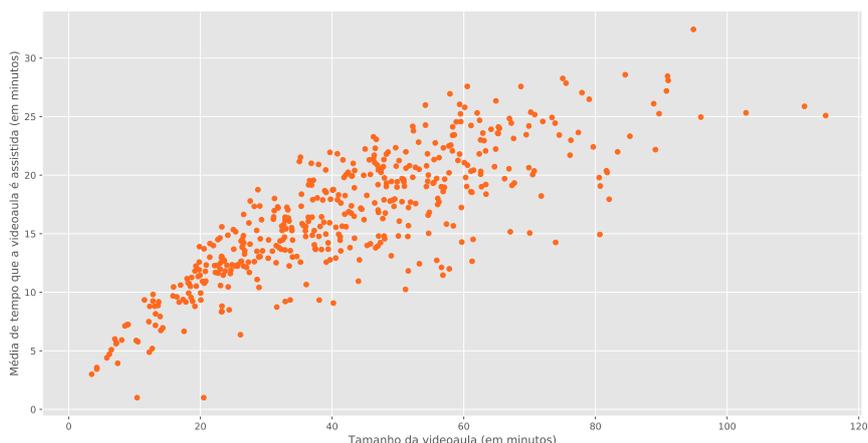


Figura 4.19: Média de tempo que cada videoaula é assistida

A tabela a seguir mostra a fração média de visualização por tempo total de vídeo. Quanto maior a duração de uma videoaula menor é a fração total da mesma, que os alunos tendem a assistir.

Pode-se observar que esses dados vão de encontro com o resultado do gráfico anterior e aplicando-se as porcentagens sobre os tempos de vídeo, tem-se no máximo 30 minutos de visualização da aula.

Esse resultado também é similar ao resultado apresentado pelo pesquisador Gaspare Bruno [15]. O seu estudo mostra que os estudantes tendem a ter baixa tolerância às aulas de longa duração (com cerca de 1h e 30 minutos de duração).

Tempo total de vídeo (em minutos)	Fração média de visualização
0 - 10	76.2 %
10 - 20	59.7 %
20 - 30	51.1 %
30 - 40	44.4 %
40 - 50	38.1 %
50 - 60	32.0 %
60 - 70	33.9 %
80 - 90	30.0 %
>90	20.6 %

Tabela 4.2: Fração média de visualização por duração da videoaula em minutos

4.7 Agrupamento por disciplinas recentes

O curso de **Sistemas em Computação** foi criado há mais de 10 anos e possui muitas videoaulas em um formato antigo e que muitas vezes pode ser pouco convidativo para determinados tipos de usuário. Além disso, outro fator que pode influenciar o número de visualizações é o conteúdo da disciplina não estar atualizado.

O objetivo dessa análise é entender se um curso mais recente possui um comportamento diferente dos cursos mais antigos.

O gráfico da figura 4.20 mostra a fração do número de visualizações pela fração do tempo de vídeo da videoaula. Nesse gráfico foram utilizadas 3725 sessões do período de 2016 a 2017. Comparado aos gráficos já apresentados, pode-se perceber que a fração de visualizações se mantém em torno de 80% do início até 90% do tempo da videoaula. Isso mostra que o interesse dos alunos se mantém quase até o final da videoaula.

Vale ressaltar que essa videoaula tem uma duração curta (**11 minutos e 35 segundos**) o que também reforça a teoria de que aulas curtas tendem a possuir melhores resultados.

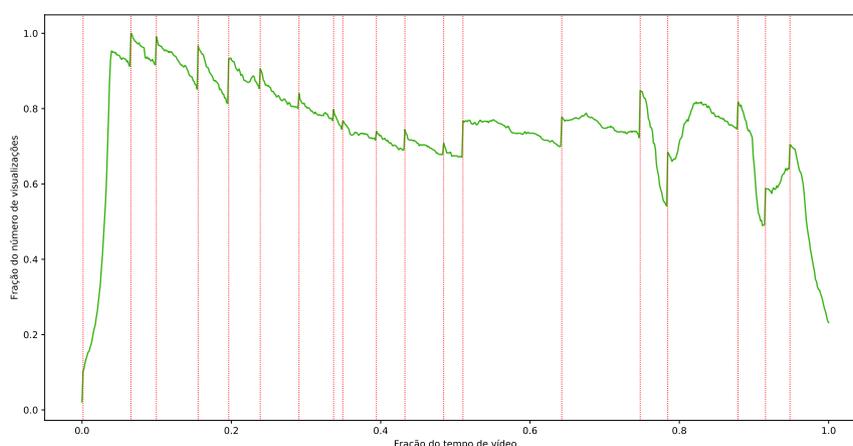


Figura 4.20: Fração do número de visualizações por fração do tempo de vídeo de uma videoaula mais recente, com duração de 11 minutos e 35 segundos

O gráfico da figura 4.21 mostra a fração do número de visualizações pela fração do tempo de vídeo da videoaula para uma aula com quase o triplo do tamanho (**30 minutos**). Nesse gráfico foram utilizadas 2327 sessões do período de 2016 a 2017. O comportamento é semelhante ao da figura 4.20

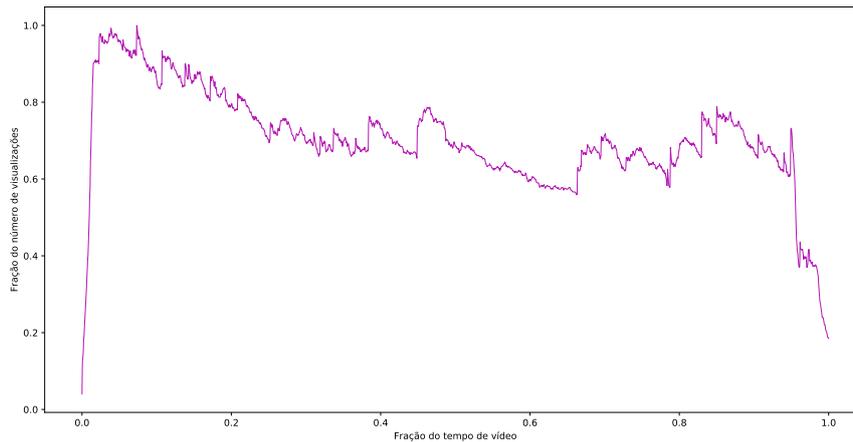


Figura 4.21: Fração do número de visualizações por fração do tempo de vídeo de uma videoaula mais recente, com duração de 30 minutos

4.8 Agrupamento por assiduidade dos cursos

Outra análise interessante a ser feita é o da assiduidade dos cursos. O objetivo é entender se os alunos se mantêm engajados ao longo do período.

A figura 4.22 apresenta dois gráficos que mostram o número de visualizações de cada videoaula do curso. O gráfico superior representa o curso **ead05020** (um curso bastante popular) e o inferior o curso **ead05029** (um curso mais recente).

Pode-se ver que o número de visualizações das últimas aulas chega a ser menor que **20%** do número das primeiras aulas. Além disso, o comportamento é bastante similar, independente do curso ser antigo ou não.

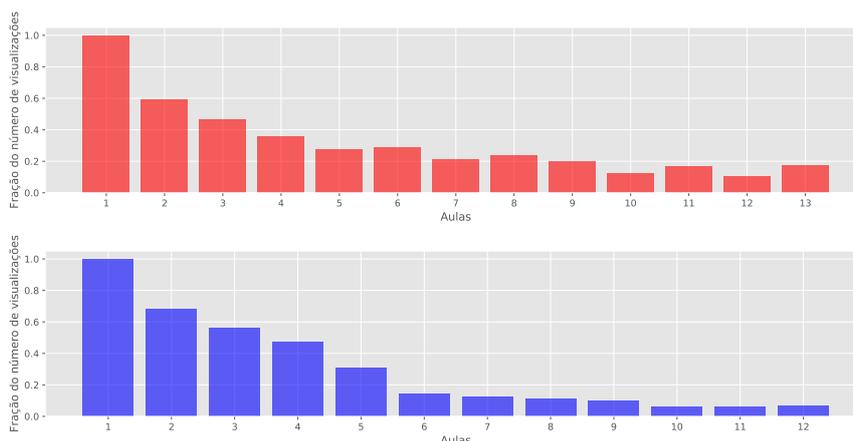


Figura 4.22: Assiduidade dos cursos

Capítulo 5

Conclusão

O trabalho aqui proposto teve uma ótima oportunidade de analisar um sistema de educação a distância, que possui milhares de alunos ao redor do Brasil.

Como exposto na introdução, as metodologias tradicionais de ensino passam por um momento de decadência e não se adequam mais aos novos tempos. Em um mundo onde a atenção é disputada por diferentes dispositivos eletrônicos e onde a informação é disseminada de maneira rápida e em grandes quantidades, torna-se cada vez mais desafiador conquistar a atenção dos alunos.

O **capítulo 4** mostrou que o sistema avaliado, possui uma amostra significativa de alunos e representa muito bem um cenário real de educação a distância.

Como visto em alguns dos gráficos, muitas aulas do curso de **Tecnologia em Sistemas de Computação** apresentam resultados pouco eficientes no que diz respeito a prender a **atenção** do aluno por um grande período de tempo. O comportamento de **navegação por tópicos** indica que cada vez mais, os alunos procuram por um **dinamismo** na aula. Não se aceita mais o **velho modelo de ensino**.

Outro resultado interessante, diz respeito a **atualização das aulas**. Aulas mais recentes apresentaram resultados melhores que aulas mais antigas. Isso pode indicar que os alunos valorizam um material renovado e que possivelmente está mais alinhado com as novas exigências do mercado ou assuntos do momento.

Também visto nas análises, os alunos tendem a assistir as aulas com **pouca repetição**, i.e, não assistem a mesma parte mais de uma vez. Esse comportamento poderia indicar que as aulas são muito eficientes e os alunos assimilam o conteúdo sem a necessidade de repetição. Por outro lado, também pode indicar a exigência pelo **dinamismo e rapidez**. Não "perde-se" mais tempo com um mesmo assunto, e o aprendizado **superficial** torna-se cada vez mais comum.

Apesar do curso avaliado no trabalho ser online e utilizar um sistema robusto de educação a distância, as aulas estudadas foram construídas baseadas no mesmo modelo de ensino das aulas presenciais. O estudo mostra que essa metodologia pode não ter sido a melhor escolha e o sistema possivelmente não atinge todo o seu

potencial.

Diante dos desafios apresentados, ficam os questionamentos de como produzir um conteúdo que consiga ser efetivo no ensino.

Como apresentado anteriormente, a participação dos alunos no processo educacional precisa ser modificada e o aluno precisa poder ser mais ativo durante a aula. Não basta apenas ouvir o professor por períodos de uma a duas horas seguidas.

- **Sistemas Adaptativos**

Um sistema mais eficiente poderia ser um que permitisse aulas adaptativas, i.e, aulas diferentes para cada tipo de aluno. Ao perceber um determinado tipo de comportamento, o sistema poderia indicar um material diferente de leitura, uma animação ou jogo interativo, ou até mesmo encurtar a aula quando percebido que o aluno absorveu a quantidade necessária para aquele momento.

- **Interação entre os alunos**

Também é preciso criar um ambiente de interação entre os estudantes. Algo pouco explorado em sistemas de EaD, são as ferramentas que permitem a comunicação e compartilhamento de informações entre os alunos de um curso. Imagine a quantidade de dúvidas, resoluções e criações produzidas por milhares de alunos ao redor do mundo. Todo esse material poderia estar documentado e ser acessado a qualquer momento por velhos e novos alunos do sistema.

- **Sistemas de recomendação**

O sistema também poderia ser capaz de fazer recomendação de conteúdo. O aluno não precisa necessariamente assistir um único curso de maneira sequencial. Existem diversos assuntos compartilhados por diferentes disciplinas.

- **Gamificação**

Uma boa estratégia para retenção dos usuários é um sistema de "gamificação" que dê "prêmios" virtuais para os usuários esforçados e ativos. No mundo das redes sociais, muitos querem compartilhar que sabem ou estão aprendendo determinado assunto e ter uma "medalha", mesmo que virtual, pode ser um diferencial.

Um projeto que pode dar continuidade ao trabalho apresentado, é a criação de uma ferramenta de dashboard que auxilie os professores a visualizarem informações sobre suas videoaulas. Esse dashboard poderia exibir gráficos e estatísticas que mostrassem com mais precisão, por exemplo, quais os pontos a serem melhorados em cada aula ou uma comparação entre as diferentes videoaulas. Isso serviria como

um termomêtro para o professor, que diante dos resultados estaria mais preparado para propor modificações e novas metodologias às suas aulas.

Espera-se que o trabalho aqui apresentado tenha mostrado a urgência de discutir os desafios da educação a distância e como o modelo tradicional precisa ser adaptado às novas tecnologias visando melhorar o desempenho dos alunos.

Referências Bibliográficas

- [1] MEC. “Educação Superior a Distância”. . <http://portal.mec.gov.br/instituicoes-credenciadas/educacao-superior-a-distancia>.
- [2] ALVES, L. “Educação a distância: conceitos e história no Brasil e no mundo”, *Revista Brasileira de Aprendizagem Aberta e a Distância*, v. 10, n. 1, 2011.
- [3] JUNIOR, J., BARROS, E., EUZÉBIO, J., et al. “Educação a distância: desafio e perspectivas”, 2015.
- [4] SOUZA E SILVA, EDMUNDO, B. G. S. D. “Máquinas Que Ensinam”. In: *Ciência para Educação, Uma ponte entre Dois Mundos*, Atheneu, cap. 12, pp. 251–268, 2017.
- [5] ABED. “Censo EAD Brasil 2016 - Relatório Analítico de Aprendizagem a Distância no Brasil”, 2016.
- [6] FREEMAN, S., EDDY, S. L., MCDONOUGH, M., et al. “Active learning increases student performance in science, engineering, and mathematics”, *Proceedings of the National Academy of Sciences*, v. 111, n. 23, pp. 8410–8415, 2014.
- [7] KRÜGER, L. M., OTHERS. “Método Tradicional e Método Construtivista de Ensino no Processo de Aprendizagem: uma investigação com os acadêmicos da disciplina Contabilidade III do curso de Ciências Contábeis da Universidade Federal de Santa Catarina”, 2013.
- [8] “Edx”. . <https://www.edx.org>. Accessed: 2018-01-18.
- [9] “Coursera”. . <https://pt.coursera.org/>. Accessed: 2018-01-18.
- [10] “Lemann”. . <http://www.fundacaolemann.org.br/coursera-brasil/>. Accessed: 2018-01-18.
- [11] “Descomplica”. . <https://www.descomplica.com.br>. Accessed: 2018-01-18.
- [12] “Videoaula@RNP”. . <https://www.videoaula.rnp.br>. Accessed: 2018-03-01.

- [13] NETTO, B. C. M., AZEVEDO, J. A., SILVA, E., et al. “Servidor Multimídia RIO em Ensino a Distância”, *6th International Free Software Fórum*, 2005.
- [14] “Entrevista — Francisco Mora, especialista em neuroeducação”. . https://brasil.elpais.com/brasil/2017/02/17/economia/1487331225_284546.html. Accessed: 2018-01-19.
- [15] BRUNO, G. G. E. “Um sistema de recomendação inteligente baseado em vídeo-aulas para educação a distância”, 2016.