



FACULDADE DE ECONOMIA E FINANÇAS IBMEC  
PROGRAMA DE PÓS-GRADUAÇÃO E PESQUISA EM  
ADMINISTRAÇÃO E ECONOMIA

DISSERTAÇÃO DE MESTRADO  
PROFISSIONALIZANTE EM ADMINISTRAÇÃO

**“MÉTRICAS DA CIRCULAÇÃO DE  
CONTÉUDOS DO MERCADO DE  
ENTRETENIMENTO NAS REDES  
P2P(Peer-to-peer): PROTÓTIPO  
EXPERIMENTAL”**

**JOÃO PAULO GINO DO REGO**

ORIENTADOR: SIMONE BACELLAR LEAL FERREIRA

Rio de Janeiro, 20 de junho de 2006

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

**“REDES P2P(PEER-TO-PEER) MÉTRICAS DA CIRCULAÇÃO DE CONTEÚDOS  
NO MERCADO DE ENTRETENIMENTO: PROTÓTIPO EXPERIMENTAL”**

JOAO PAULO GINO DO REGO

Dissertação apresentada ao curso de  
Mestrado Profissionalizante em  
Administração como requisito parcial para  
obtenção do Grau de Mestre em  
Administração.  
Área de Concentração: Sistemas de  
Informação

ORIENTADOR: SIMONE BACELLAR LEAL FERREIRA

Rio de Janeiro, 20 de junho de 2006.



**“REDES P2P(PEER-TO-PEER) MÉTRICAS DA CIRCULAÇÃO DE CONTEÚDOS  
DO MERCADO DE ENTRETENIMENTO: PROTÓTIPO EXPERIMENTAL”**

JOÃO PAULO GINO DO REGO

Dissertação apresentada ao curso de  
Mestrado Profissionalizante em  
Administração como requisito parcial para  
obtenção do Grau de Mestre em  
Administração.  
Área de Concentração: Sistemas de  
Informação

Avaliação:

BANCA EXAMINADORA:

---

Doutora SIMONE BACELLAR LEAL FERREIRA (Orientador)  
Instituição: IBMEC

---

Doutora MARIA AUGUSTA SOARES MACHADO  
Instituição: IBMEC

---

Doutora MARIE AGNES CHAUVEL  
Instituição: IAG / PUC

Rio de Janeiro, 20 de junho de 2006.

FICHA CATALOGRÁFICA

Entrar em contato com a biblioteca no 14º andar,  
ou através do e-mail: [geysa@ibmecrj.br](mailto:geysa@ibmecrj.br)

## **RESUMO**

Este estudo visa demonstrar através de um protótipo computacional, a geração de estatísticas, sobre a circulação de conteúdos específicos, dentro de uma rede P2P. A indústria do entretenimento, que detém direitos sobre os conteúdos, tais como músicas, filmes e livros vem tendo grandes perdas financeiras, com a circulação gratuita através da Internet. Este trabalho analisa as topologias de redes P2P existentes e suas variâncias, assim como a evolução dos meios de digitalização numa linha cronológica. Somado a esses estudos, uma breve análise dos fatores sócio-culturais envolvidos, na migração do conceito de propriedade para o mundo digital.

Palavras Chave:

P2P, Redes, fonográfico, cinematográfico, edonkey, propriedade, digital.

## LISTA DE FIGURAS

Figura 1 - Suportes de Reprodução de Áudio (LP,K7,CD).....	13
Figura 2 - Suportes de Vídeo (VHS, BetaMax, DVD).....	15
Figura 3 - Discos HD DVD e BlueRay .....	16
Figure 4 - Onda analógica .....	17
Figure 5 - Onda digitalizada .....	17
Figura 6 - Sinais digitais físicos num CD.....	18
Figura 7 - Fluxo de Compressão / Transporte / Descompressão de Dados.....	19
Figura 8 - Fonte primária de dados.....	20
Figura 9 - Fonte primária em pixels (64x64).....	20
Figura 10 - Representação alfabética da matriz.....	21
Figura 11 - Demonstração de processo de compressão / descompressão.....	22
Figura 12 - Topologia Centralizada.....	31
Figura 13 - Topologia em Anel .....	32
Figure 14 - Topologia Hierárquica .....	33
Figura 15 - Topologia Descentralizada .....	33
Figura 16 - Topologia Centralizada + Topologia Anel .....	34
Figura 17 - Topologia Centralizada + Topologia Centralizada.....	35
Figura 18 - Topologia Centralizada + Topologia Descentralizada .....	35
Figura 19 - Fluxo metodologia XP .....	45
Figura 20 - Tela do Site BigChampane .....	48
Figura 21 - Exemplo do Código do Protótipo em Python .....	50
Figura 22 - Tela do Cliente Emule .....	52
Figure 23 - Evento de Login com o servidor.....	53
Figura 24 - Tela do software Ethereum .....	54
Figura 25 - Representação de um Grafo.....	56
Figure 26 - Relacionamento dos Nós de uma rede P2P .....	57
Figura 27 - Tela do protótipo em funcionamento.....	59
Figure 28 - Lista de arquivos com seus respectivos FILE ID .....	60
Figura 29 – Usuários únicos com o filme “Scary Movie 4”.....	62
Figura 30 – Caixa do Filme “Scary Movie 4” .....	63
Figura 31 - Estimativa de perda financeira.....	<b>Erro! Indicador não definido.</b>

## LISTA DE TABELAS

Tabela 1 - Comparativo de suportes de reprodução de áudio .....	14
Tabela 2 - Comparativo de suportes de reprodução de vídeo .....	15
Tabela 3 – Sistemas Distribuídos – Características .....	30
Tabela 4 - Definições P2P / Autores .....	36
Tabela 5 - Nomenclaturas de Schollmeie .....	39
Tabela 6 - Horas gastas para produção do protótipo .....	58
Tabela 7 - Testes de Conteúdo .....	60
Tabela 8 – Outros Dados .....	64

## SUMÁRIO

<b>RESUMO</b> .....	<b>1</b>
<b>LISTA DE FIGURAS</b> .....	<b>2</b>
<b>LISTA DE TABELAS</b> .....	<b>3</b>
<b>SUMÁRIO</b> .....	<b>4</b>
<b>1 INTRODUÇÃO</b> .....	<b>6</b>
<b>1.1 FORMULAÇÃO DO PROBLEMA</b> .....	<b>8</b>
<b>1.2 PERGUNTA DE TESE</b> .....	<b>8</b>
<b>1.3 OBJETIVO PRINCIPAL</b> .....	<b>9</b>
<b>1.4 OBJETIVO FINAL</b> .....	<b>9</b>
<b>1.5 OBJETIVOS INTERMEDIÁRIOS</b> .....	<b>9</b>
<b>1.6 JUSTIFICATIVA E RELEVÂNCIA DA PESQUISA</b> .....	<b>10</b>
<b>1.7 LIMITAÇÃO DA PESQUISA</b> .....	<b>10</b>
<b>2 REREFENCIAL TEÓRICO I</b> .....	<b>12</b>
<b>2.1 O ENTRETENIMENTO E A DIGITALIZAÇÃO</b> .....	<b>12</b>
<b>2.2 EVOLUCAÇÃO DA INDÚSTRIA DO ENTRETENIMENTO</b> .....	<b>13</b>
2.2.1 <i>MERCADO FONOGRÁFICO</i> .....	<b>13</b>
2.2.2 <i>MERCADO CINEMATOGRAFICO</i> .....	<b>14</b>
<b>2.3 DA COMUNICAÇÃO ELÉTRICA, DIGITALIZAÇÃO</b> .....	<b>16</b>
2.3.1 <i>DIGITALIZAÇÃO</i> .....	<b>16</b>
2.3.2 <i>COMPRESSÃO DE DADOS</i> .....	<b>19</b>
<b>2.4 FORMATOS DIGITAIS DE ÁUDIO E VÍDEO</b> .....	<b>23</b>
2.4.1 <i>MP3</i> .....	<b>23</b>
2.4.2 <i>MPEG-4</i> .....	<b>25</b>
<b>3 REREFENCIAL TEÓRICO II</b> .....	<b>27</b>
<b>3.1 DESAFIOS DA INDÚSTRIA DO ENTRETENIMENTO</b> .....	<b>27</b>
<b>3.2 REDE DE COMPUTADORES</b> .....	<b>28</b>
<b>3.3 SISTEMAS DISTRIBUÍDOS</b> .....	<b>29</b>
3.3.1 <i>TOPOLOGIAS DE SISTEMAS DISTRIBUÍDOS</i> .....	<b>30</b>
3.3.2 <i>TOPOLOGIAS DE SISTEMA SISTEMAS DISTRIBUÍDOS PURAS</i> .....	<b>31</b>
3.3.2.1 <i>TOPOLOGIA CENTRALIZADA</i> .....	<b>31</b>
3.3.2.2 <i>TOLOGIA EM ANEL</i> .....	<b>32</b>
3.3.2.3 <i>TOLOGIA HIERÁRQUICA</i> .....	<b>32</b>
3.3.2.4 <i>TOLOGIA DESCENTRALIZADA</i> .....	<b>33</b>
3.3.3 <i>TOPOLOGIA DE SISTEMAS DISTRIBUÍDOS HÍBRIDAS</i> .....	<b>34</b>
3.3.3.1 <i>TOLOGIA CENTRALIZADA + TOPOLOGIA ANEL</i> .....	<b>34</b>
3.3.3.2 <i>TOLOGIA CENTRALIZADA + TOPOLOGIA CENTRALIZADA</i> .....	<b>34</b>
3.3.3.3 <i>TOLOGIA CENTRALIZADA + TOPOLOGIA DESCENTRALIZADA</i> .....	<b>35</b>
<b>3.4 REDES P2P</b> .....	<b>36</b>
3.4.1 <i>DEFINIÇÕES DE REDES P2P</i> .....	<b>36</b>
3.4.2 <i>FUNCIONAMENTO DAS REDES P2P</i> .....	<b>38</b>
3.4.3 <i>NOMECLATURAS DAS REDES P2P</i> .....	<b>38</b>

3.4.4	REDE P2P PURA.....	39
3.4.5	REDE P2P HÍBRIDA .....	39
<b>3.5</b>	<b>PROCOLOS P2P .....</b>	<b>40</b>
3.5.1	PROCOLO FASTTRACK.....	40
3.5.2	PROCOLO EDONKEY.....	41
3.5.3	PROCOLO GNUTELLA.....	42
3.5.4	PROCOLO BITTORRENT .....	42
3.5.5	PROCOLO FREENET.....	43
<b>4</b>	<b>METODOLOGIA .....</b>	<b>44</b>
<b>4.1</b>	<b>MOTIVAÇÃO .....</b>	<b>44</b>
<b>4.2</b>	<b>METODOLOGIA DE DESENVOLVIMENTO DO PROTÓTIPO .....</b>	<b>44</b>
<b>4.3</b>	<b>REQUISITOS DO PROTÓTIPO.....</b>	<b>46</b>
<b>4.4</b>	<b>CONCEPÇÃO DO PROTÓTIPO.....</b>	<b>47</b>
<b>4.5</b>	<b>PROJETOS ANTERIORES.....</b>	<b>47</b>
<b>4.6</b>	<b>O PROTÓTIPO .....</b>	<b>48</b>
4.6.1	LINGUAGEM DE PROGRAMAÇÃO.....	48
4.6.2	LINGUAGEM PYTHON.....	49
4.6.3	BIBLIOTECA TWISTED .....	50
4.6.4	REFERÊNCIAS DO PROCOLO EDONKEY .....	51
4.6.5	ESTRUTURA DO PROCOLO.....	53
4.6.6	CRIAÇÃO DE AMBIENTE DE TESTES .....	54
4.6.7	PORTAS E CONEXÕES.....	55
4.6.8	GERAÇÃO DE ESTATÍSTICAS .....	55
4.6.9	BUSCA EM PROFUNDIDADE .....	55
<b>4.7</b>	<b>CRONOGRAMA DO PROTÓTIPO .....</b>	<b>57</b>
4.7.1	AUXÍLIO EXTERNO .....	58
<b>5</b>	<b>ANALISE DOS RESULTADOS.....</b>	<b>59</b>
<b>5.1</b>	<b>PERÍODO EXPERIMENTAL .....</b>	<b>59</b>
<b>5.2</b>	<b>SEQÜÊNCIA DE USO DO PROTÓTIPO: .....</b>	<b>60</b>
<b>5.3</b>	<b>ANÁLISES EFETUADAS.....</b>	<b>60</b>
<b>5.4</b>	<b>ANÁLISE DOS RESULTADOS.....</b>	<b>61</b>
<b>5.5</b>	<b>ESTIMATIVA DE INTERESSE DO PÚBLICO .....</b>	<b>61</b>
<b>5.6</b>	<b>ESTIMATIVA DE PERDAS FINANCEIRAS: .....</b>	<b>63</b>
<b>5.7</b>	<b>OUTROS DADOS PERTINENTES LEVANTADOS .....</b>	<b>64</b>
<b>6</b>	<b>CONCLUSÕES.....</b>	<b>65</b>
<b>6.1</b>	<b>RECOMENDAÇÃO PARA TRABALHOS FUTUROS.....</b>	<b>67</b>
6.1.1	PROTÓTIPO:.....	67
6.1.2	PESQUISAS.....	67
<b>7</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>69</b>

## 1 INTRODUÇÃO

*“O que poderia ser mais confuso do que um mundo que não depende de ninguém e que a todos influencia”* (Doc Searls, 2000).

A economia está sempre em transformação, e a essência do conceito de propriedade, como conhecemos, sofreu alterações. Ela perdeu seu valor, e está sendo substituída. Na dinâmica global as pessoas passaram a valorizar mais, experiências menos duradouras, que não necessariamente, se materializem em bens. [Rifikin, 2000]

Nesse contexto, a Internet favoreceu a formação de sociedades digitais (virtuais), e enfraquece aos poucos o antigo sistema capitalista, no qual as antigas leis econômicas prevaleceram. [Rifikin, 2000] [Hedlund, 2003]

A Internet contribuiu diretamente para a aceleração deste novo modelo sócio-cultural. Fazendo com que conteúdos (áudio, vídeos, textos) digitalizados circulassem pela grande rede. Permitindo então, a quebra do paradigma do suporte (meio qual o conteúdo é distribuído), fazendo com que ele, não tivesse mais a importância de outrora, formando os novos pilares comportamentais. [Rifikin, 2002]

Embalados por essas mudanças, o público da Internet, essencialmente jovem e culturalmente avançado, tem usufruído dos conteúdos digitais, passando horas conectados, navegando, em busca desta nova forma de entretenimento. [Tapscott, 1999]

Com a chegada da banda larga, se tornou possível que dados fluíssem em alta velocidade, dando suporte a essa nova e crescente demanda de conteúdos mais complexos (vídeos). As

redes de banda larga já se espalharam pelo mundo; nos EUA, por exemplo, já representam mais de 40% dos lares americanos. No Brasil, nos últimos dois anos, a oferta cresceu, o custo diminuiu, existindo hoje mais de três milhões de domicílios conectados a altas velocidades [Ibope//NetRatings].

Durante o próprio amadurecimento da Internet, surgiram novas redes dentro dela mesma, chamadas P2P (peer-to-peer), que permitiam o compartilhamento ou troca de todo tipo (áudio, vídeo, texto) de conteúdo de forma facilitada e colaborativa. Resumidamente, uma rede P2P é composta por dois elementos, os nós e o protocolo. Os nós (ou peers), são os dispositivos (computadores) que compõem o a malha da rede e detém o conteúdo; e o protocolo, é o idioma utilizado pelos nós para se comunicarem. [Figueiredo, 2006]

Com a evolução e o desenvolvimento das redes P2P, arquivos com os mais diversos conteúdos circulam livremente pela Internet, sem quase nenhum tipo de controle ou autorização de seus autores para tal. Com isso, a indústria do entretenimento, que muitas vezes se beneficiou da natureza consumista do jovem, passou a sofrer as conseqüências da desta nova geração de consumidores digitais.

O setor de entretenimento foi o setor que mais tem tido perdas financeiras diretas, com a circulação não autorizada de material protegido por direitos de propriedade autoral, como músicas, livros, filmes, etc., nas redes P2P[http\_5].

Nesta geração digital, a sociedade é em rede, e nada pertence a um só indivíduo; pertence sempre a comunidade, que socializa seus conteúdos e cria suas próprias regras de conduta, alheia às leis do mundo real (físico). [Tapscott, 1999]

Com o crescimento das comunidades P2P, os tribunais do mundo todo começaram a abrir processos contra programadores dos protocolos e proprietários de servidores das redes P2P. Até mesmo usuários que, de alguma forma contribuíssem para a facilitação ou divulgação não autorizada de conteúdos protegidos intelectualmente.[http\_7]

Independentemente de essas medidas, as comunidades P2P cresceram, se sofisticaram, e conseguem um número cada vez maior de adeptos.[http\_7]

Motivado por esse crescimento das redes P2P, o presente trabalho visa desenvolver um protótipo computacional que colete dados pertinentes para a indústria do entretenimento, que permita aos administradores quantificar o tráfego de seus conteúdos através das redes P2P.

## **1.1 FORMULAÇÃO DO PROBLEMA**

Atualmente tem-se observado a dificuldade dos profissionais de gestão e marketing em coletar dados relativos à circulação de conteúdos específicos na Internet nas redes de troca de arquivos. Como na pesquisa realizada para este trabalho não foi encontrado nenhum software ou qualquer outro mecanismo que auxiliasse na tarefa de distinguir conteúdos e prover dados analíticos sobre o tráfego dos mesmos, resolveu-se abordar esse assunto e desenvolver um protótipo computacional para este fim.

## **1.2 PERGUNTA DE PESQUISA**

É possível desenvolver um protótipo computacional capaz de coletar informações sobre a circulação de conteúdos específicos, nas redes P2P?

### **1.3 OBJETIVO PRINCIPAL**

O principal objetivo do presente trabalho é desenvolver um protótipo capaz de coletar informações sobre conteúdos específicos que circulam através das redes P2P.

### **1.4 OBJETIVO FINAL**

O objetivo final da presente pesquisa é de prover subsídios estatísticos para profissionais de marketing e gestão, sobre a circulação de conteúdos distintos nas redes P2P.

### **1.5 OBJETIVOS INTERMEDIÁRIOS**

- a) Realizar estudo teórico sobre as principais redes P2P existentes até o término deste trabalho.
- b) Analisar sua estrutura e tabular suas diferenças de forma comparativa, tendo como parâmetro seu funcionamento.
- c) Analisar quais as redes P2P dão maior facilidade para se distinguir os conteúdos circulados.
- d) Criar protótipo que instrumentalize o objetivo da pesquisa.
- e) Utilizar o protótipo para avaliar um(1) conteúdo distinto de cada origem ( Nacional / Internacional), por tipo de conteúdo:
  - Faixa de Áudio (CD)
  - Filme Completo (DVD)
  - Texto Completo (Livro)
- f) Analisar um resultado obtido nos casos de uso propostos.

- g) Auxiliar o processo de tomada de decisão em empresas do segmento de entretenimento, com base nos subsídios estáticos fornecidos pelo protótipo.

## **1.6 JUSTIFICATIVA E RELEVÂNCIA DA PESQUISA**

A pesquisa tal como o desenvolvimento da ferramenta se faz necessária pela ausência de um software até o término deste trabalho com as características abaixo:

- Busca de Música por (Obra / Autor / Coletânea), Filmes ou Livros
- Mapeamento do tráfego de um conteúdo em um intervalo de tempo
- Capacidade de computar o maior número de usuários únicos durante o período.
- Seleção de um, ou mais (servidores / “trackers” / nós raízes).
- Saída em formato XML

Com essas características, o protótipo pode coletar dados sobre a circulação de um conteúdo específico, em um intervalo de tempo, numa das redes P2P existentes.

## **1.7 DELIMITAÇÃO DA PESQUISA**

- Os dados fornecidos pelo protótipo não registram individualmente os usuários das redes P2P. Ou seja, não possibilita a identificação pessoal dos usuários.
- Não será analisada profundamente a questão ética ou legal sobre a existência das redes P2P.
- A pesquisa não tem a pretensão de esgotar a leitura das redes P2P ou ser uma ferramenta definitiva para estudos neste vasto campo.
- Os casos de testes apresentados foram realizados com apenas três tipos de conteúdos (áudio, vídeo, e textos), não sendo realizados testes com conteúdos de outra natureza.



## **2 REREFENCIAL TEÓRICO I**

### **2.1 O ENTRETENIMENTO E A DIGITALIZAÇÃO**

O aumento da utilização das novas tecnologias nos meios de comunicação tem derrubado as fronteiras geográficas, facilitando dessa forma a agilidade no fluxo de informações pelo mundo. É possível escutar uma rádio do Canadá e ao mesmo tempo acessar um site de eventos esportivos japoneses, independentemente do lugar do planeta em que o usuário está conectado. Através da Internet e dos acessos em alta velocidade esta conectividade tem elevado os conceitos da comunicação a patamares jamais vistos, tornando o entretenimento um estilo de vida, e transformando o lúdico em consumo. [Castells, 2003]

A conectividade trouxe novo impulso econômico e novas possibilidades para os mercados de entretenimento em massa. As rádios já vinham sofrendo quedas de audiência, desde o surgimento da televisão, em meados dos anos 70. Entretanto, por conta das pobres grades de programação e pela falta de interatividade com os telespectadores, a televisão também não prende mais a atenção do público como antigamente. Neste ínterim, a Internet surgiu como um meio capaz de integrar texto, áudio e vídeo, simultaneamente, de modo a complementar os meios de comunicação existentes até então. [Noam e Green,2003]

Por esses motivos, a Internet está se consolidando como um meio de comunicação, e explorá-la de forma comercial tem se tornado essencial para a sobrevivência da indústria do entretenimento, já que as empresas deste segmento têm, atualmente, pouco ou nenhum controle sobre o conteúdo difundido na rede mundial de computadores. [Bandeira, 2004]

Para compreender a evolução da indústria e os objetos técnicos da digitalização é necessário para o bom entendimento do cenário que este estudo aborda.

## 2.2 EVOLUCAÇÃO DA INDÚSTRIA DO ENTRETENIMENTO

### 2.2.1 MERCADO FONOGRÁFICO

Na maioria das indústrias, a dinamização do parque tecnológico é uma constante. Não é diferente na indústria do entretenimento, onde as mudanças dos suportes de reprodução sonora se tornaram marcos. O mercado fonográfico, por exemplo, vivenciou a substituição em massa do antigo *LP (Long-Play)* pelo moderno *CD (Compact-Disc)* no início de 1988. (Marchi, L. 2005) Vários suportes intermediários ou alternativos surgiram nesse ínterim, como o *K7* e o *Minidisc*, que não alcançaram à mesma popularidade. Num segundo momento, bem mais curto, desta evolução tecnológica, surgiu o *DVD-Audio* prometendo mais fidelidade sonora e opções acústicas de reprodução. [http\_8] A figura 1 mostra alguns destes suportes de áudios em massa.



**Figura 1 - Suportes de Reprodução de Áudio (LP,K7,CD)**

	Medidas	Manuseio	Qualidade	Leitura	Capacidade	Mercado
LP	31 cm (diâmetro)	Difícil	Boa	Mecânica	20 minutos (face)	1950
K7	10x20 cm	Simples	Média	Mecânica	30 minutos (face)	1985

Mini-Disc	10x5 cm	Simples	Ótima	Laser	60 minutos	1992
CD	10 cm (diâmetro)	Médio	Ótima	Laser	72 minutos	1988
DVD Áudio	10 cm (diâmetro)	Médio	Excelente	Laser	240 minutos	1999

**Tabela 1** - Comparativo de suportes de reprodução de Áudio.

Durante os processos de migração de suporte, as medidas, facilidade no manuseio, qualidade, tipo de leitura utilizado e a minutagem variaram, como demonstrado na tabela 1.

### 2.2.2 MERCADO CINEMATOGRAFICO

Assim como o mercado fonográfico, o cinematográfico passou pelo mesmo processo de evolução tecnológica, dos rolos de filmes até os *DVDs*. Porém, de forma mais lenta, por questões políticas nos consórcios das empresas desenvolvedoras [http\_9].

Enquanto as empresas atuantes no mercado musical já colhiam os frutos da migração para um suporte digital (CD), o mercado de vídeo se conflitava politicamente para a escolha de um padrão de vídeo digital. Um consórcio liderado pela *Philips* e *Sony* defendia o uso do *MultiMedia Compact Disc* (MMCD), e o seu concorrente direto, o *Super Density Disc* (SD), era patrocinado pela *Toshiba*, *Time-Warner*, *Matsushita Electric*, *Hitachi*, *Mitsubishi*, *Pioneer*, *Thomson*, e *JVC*. Para não ocorrer disputa de padrões no mercado, como no caso do *VHS*, que concorria com o padrão *BetaMax*, o então presidente da *IBM*, *Lou Gerstner*, propôs a união dos padrões propostos através da junção das melhores especificações que cada um oferecia. *Philips* e *Sony* abandonaram sua proposta adicionando ao padrão SD a capacidade de pular faixas e a especificação de dados capaz de tornar o suporte (disco) tolerante a falhas (pequenos arranhões) [http\_9]. A figura 2 mostra a imagem de alguns dos suportes de vídeo em massa.



**Figura 2 -** Suportes de Vídeo (VHS, BetaMax, DVD)

Em 1995, surgiu a especificação do DVD, e um ano depois, as primeiras unidades foram produzidas no Japão. O preço desta tecnologia foi caindo com o passar dos anos e sua base de usuários se consolidou. No Brasil, o DVD só se tornou popular em 2003, com o grande número de lançamentos a preços populares. A popularidade do padrão de DVD praticamente extinguiu lançamentos de filmes em *VHS* em 2001. [http\_10]



**Tabela 2 -** Comparativo de suportes de reprodução de vídeo

Nesta migração de suporte, assim como no caso do áudio, variaram as medidas do suporte, sua facilidade de manuseio, qualidade, tipo de leitura utilizado e minutagem suportados, como demonstrado na tabela 2.

Para compensar o atraso do lançamento do DVD, agora o mercado cinematográfico já prepara o sucessor do DVD. Estão no páreo os formatos Blue-Ray e o HD DVD. Ambos com capacidade acima de 30 Gigabytes por disco e imagens com o dobro da resolução dos atuais DVDs.[http\_11] A figura 3 mostra as imagens dos possíveis sucessores do DVD.



**Figura 3 - Discos HD DVD e BlueRay**

## **2.3 DA COMUNICAÇÃO ELÉTRICA, DIGITALIZAÇÃO**

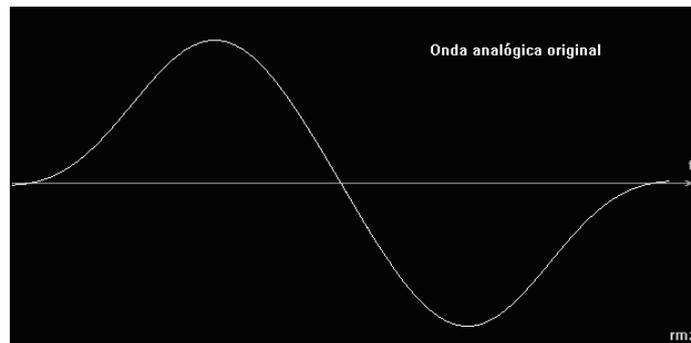
No mundo da evolução tecnológica a ampliação das capacidades de comunicação sempre foi um fator motivador para o homem. Tentando mudar a percepção da comunicação um sueco chamado *Samuel Morse* (1837) inventou o telégrafo elétrico. Sua intenção era que uma conversa pudesse se realizar sem que os participantes estivessem presentes no mesmo local. Ele codificou 30 sinais seqüenciais com dois símbolos, um ponto e um traço, para representar todo o alfabeto. O sinal do traço era executado com a continuidade do sinal elétrico por um período superior a um segundo, apertando-se uma haste de metal. Já para representar o ponto, bastava que a pressão da haste de metal fosse próxima a metade do período do traço. Através desta representação simples era possível realizar a transferência de mensagens completas, de um ponto a outro. Esta invenção foi a primeira a processar eletricamente informações para fins de comunicação. [Santos, M. 2002]

### **2.3.1 DIGITALIZAÇÃO**

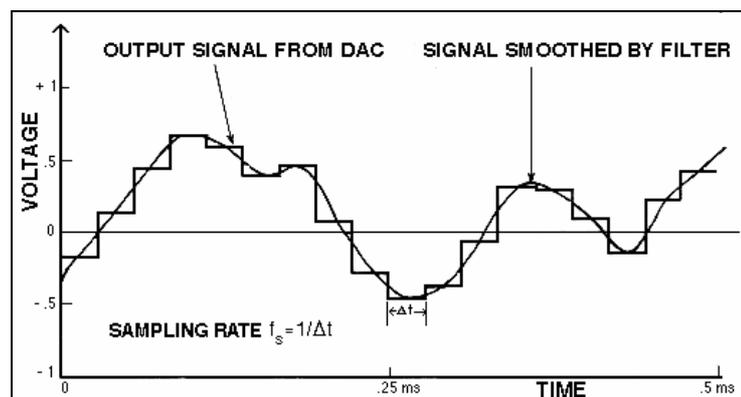
Da mesma forma, os computadores desde os seus primórdios até hoje, compreendem uma representação de dois estágios, sendo que ao invés de pontos e traços como nos antigos telégrafos, as informações são representadas de forma binária (zero e um). Essas estruturas permitiram que todo tipo de informação pudesse ser representado, e informações complexas como textos, vídeos e músicas pudessem também ser codificados.

Para entender como um conteúdo complexo pode ser eletronicamente representado é necessário analisar o campo da digitalização que permite a transformação de informações analógicas para digitais.

Um exemplo clássico na digitalização de informações é a digitalização de ondas sonoras. Na figura 4 está representada graficamente uma onda sonora demonstrada analogicamente. Suas curvas representam à amplitude de cada onda, e a seqüência delas representa a freqüência.



**Figura 4 - Onda analógica**



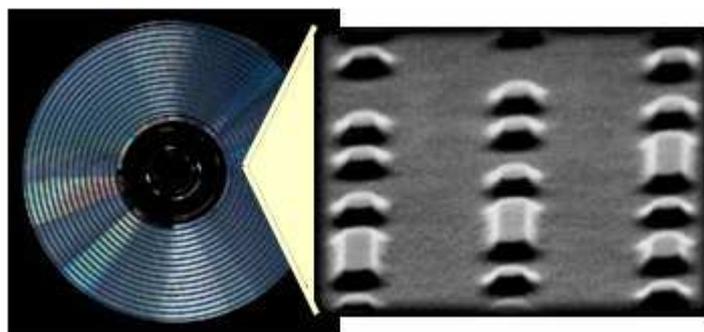
**Figura 5 - Onda digitalizada**

Observando a figura 5, pode-se perceber que as ondas estão representadas de outra forma, se comparadas à figura anterior. Elas estão representadas de forma digital, com escalas de zeros

e uns de maneira a desenhar o contorno da sua respectiva analógica. Quanto menor a distância entre os degraus de sinais digital, mais fiel à onda analógica será de sua equivalente digital.

De igual forma a mesma teoria é aplicada a conteúdos multimídia (áudio + vídeo). Diferindo apenas na adição de uma série de sinais para compor cada quadro da imagem, o som, além das definições de luminância, contraste e brilho. Todas essas informações são unidas (multiplexadas) numa única onda, formando sua equivalente digital de informação. Um bom exemplo do uso de informação digitalizada é o CD, o Compact Disc, como já visto anteriormente introduzido no mercado em 1989. [Dias, M. 2000].

A figura 6 apresenta os sulcos de um CD prensado por processo industrial em disco de acrílico com camada reflexiva e uma fina película de policarboneto com seus sinais marcados.



**Figura 6 - Sinais digitais físicos num CD**

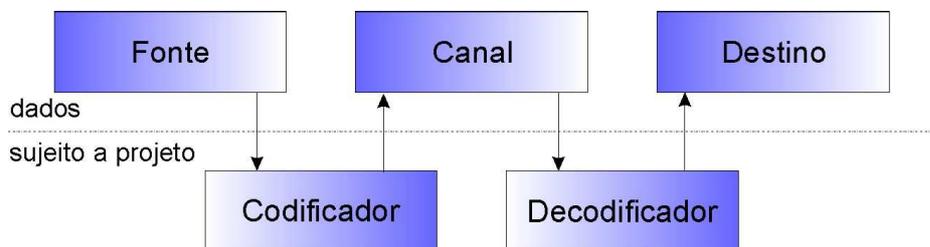
Os sinais digitais estão representados no CD de forma binária. Os traços longos representam o algarismo 1 (um) e os curtos o 0 (zero). Através dessas informações um tocador de CD utiliza um feixe de laser capaz de perceber a diferença destes sinais através dos sulcos da superfície, transportando as informações adquiridas através de um chip especialista que decodifica os

sinais e traduzindo-os numa seqüência de sinais elétricos nas caixas acústicas gerando frequências sonoras.

Com o entendimento acerca dos processos de digitalização, é necessário compreender também os formatos de compressão de dados utilizados nos arquivos utilizados pelas redes P2P.

### 2.3.2 COMPRESSÃO DE DADOS

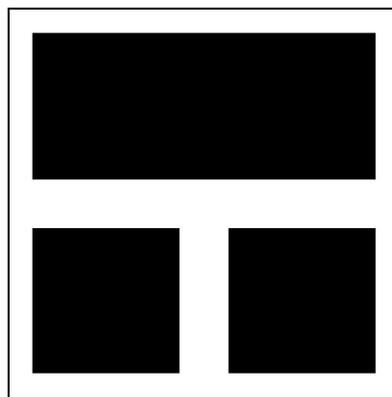
A compressão de dados é um campo da matemática para computação muito vasta, que visa diminuir (comprimir) um determinado conteúdo. Neste estudo será demonstrado um processo completo de compressão, transporte e descompressão de uma fonte primária de dados. Este processo poder ser visto na figura 7.



**Figura 7** - Fluxo de Compressão / Transporte / Descompressão de Dados

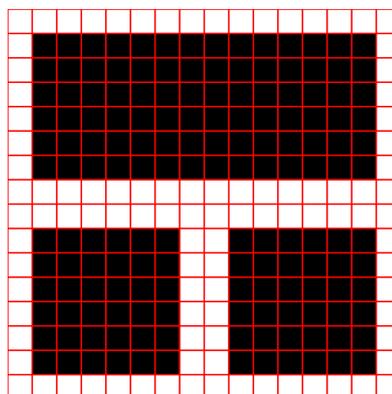
O modelo da figura 7 deixa expresso, que uma fonte de dados primária que passa pelo codificador terá seu conteúdo comprimido. Desta forma, a fonte já codificada segue pelo canal de transporte com menos volume e chega ao decodificador para ser descomprimida, voltando ao seu estado original. Este processo não altera o conteúdo.

Para exemplificar como esse processo acontece na prática será realizada uma demonstração utilizando como fonte primária de dados a figura 8.



**Figura 8** - Fonte primária de dados

A fonte primária de dados precisa estar representada de alguma forma digital(discreta) para que possa ser codificada. Para visualizar melhor a maneira como isto é realizado, a fonte primária de dados foi dividida em quadrantes. E a cada unidade criada desta delegáramos o nome de pixel. Que é a menor unidade que compõem uma imagem digital.



**Figura 9** - Fonte primária em pixels (64x64)

Como se observa na figura 10 percebe-se uma imagem com 64 pixels de altura por 64 de comprimento. Além de se enxergar com mais clareza as divisões existentes entre cada polígono dentro da figura. Uma vez tendo a figura representada de uma maneira digital, pode-se iniciar o projeto de codificação matriz da imagem, que será realizada com dois símbolos, A letra “A” e “B”. A cor branca será representada por “A” e a preta pela letra “B”.

A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A
A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A
A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A
A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A
A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A
A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	B	B	B	B	B	B	A	A	B	B	B	B	B	B	A
A	B	B	B	B	B	B	A	A	B	B	B	B	B	B	A
A	B	B	B	B	B	B	A	A	B	B	B	B	B	B	A
A	B	B	B	B	B	B	A	A	B	B	B	B	B	B	A
A	B	B	B	B	B	B	A	A	B	B	B	B	B	B	A
A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A

**Figura 10** - Representação alfabética da matriz

Uma vez tendo esta representação alfabética da imagem pode-se iniciar o processo de compressão da mesma. Para fins didáticos, será aplicada uma compressão simplista que teve sua criação para uso em aparelhos de telemetria. A compressão RLE (*Run-length encoding*) é uma técnica muito popular, por ser simples de se implantar e possuir bons resultados. Muitos formatos de dados se basearam nas suas idéias, nos primórdios da computação gráfica, e até hoje está presente em diversas aplicações e aparelhos como o de fax. [KAU, 1995]

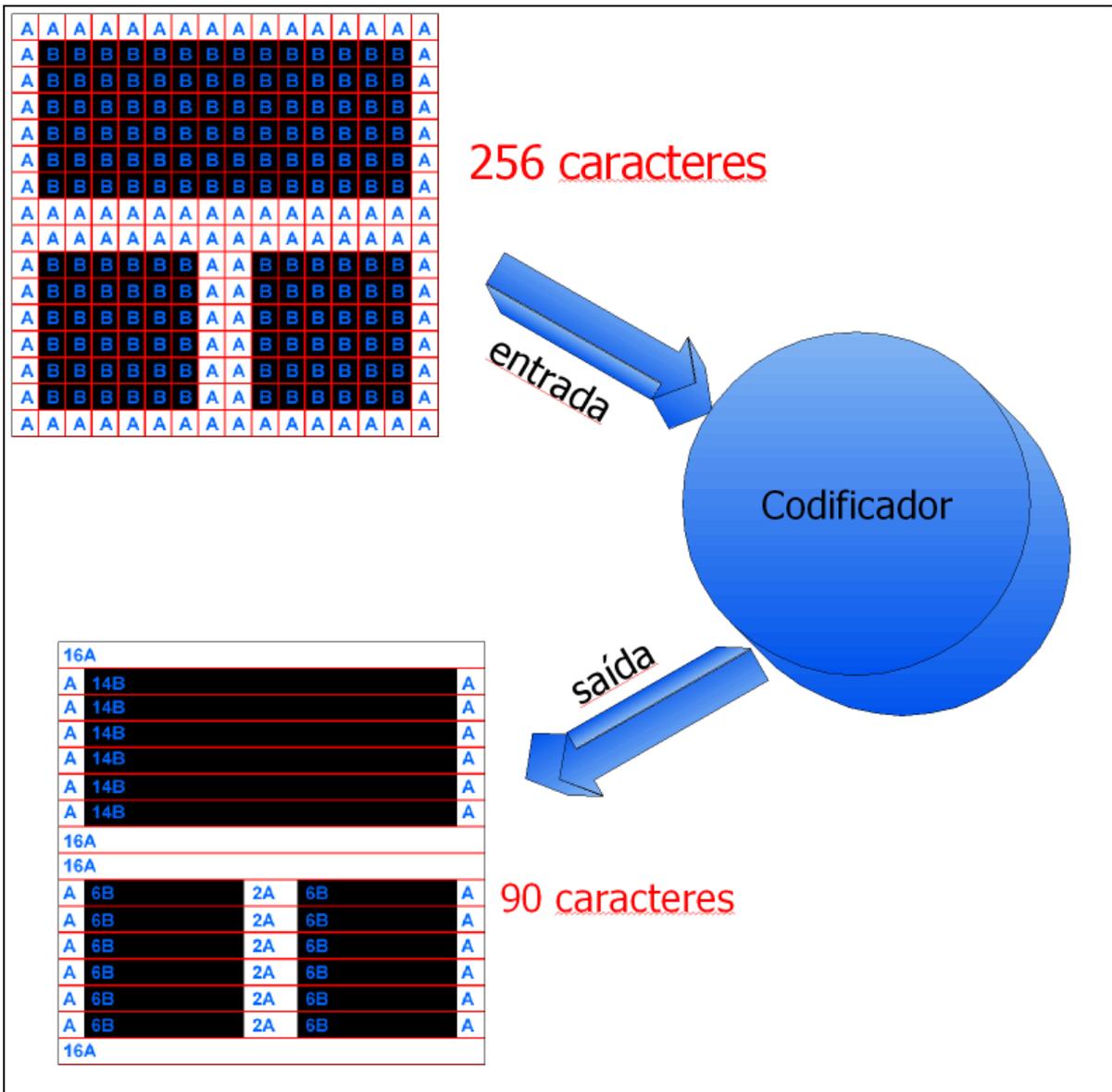


Figura 11 - Demonstração de processo de compressão

Já para descomprimir a informação o decodificador agiria da seguinte maneira:

**Decodificador** = Repetir 16 vezes “A”;  
 “A”, Repetir 14 vezes “B”, “A”;  
 Etc...

Como se pode observar na demonstração, é perfeitamente possível “diminuir” uma informação e garantir sua integridade sem perdas. Técnicas similares a essa, porém bem mais

complexas são adotadas para conteúdos de áudio e vídeo. Estas técnicas resultaram em formatos digitais, que por sua vez, contribuíram diretamente para o surgimento das redes P2P.

## **2.4 FORMATOS DIGITAIS DE ÁUDIO E VÍDEO**

A música sempre exerceu grande influência sobre o homem. Por isso, o mercado fonográfico soube retirar dela bons lucros, transformando-a em mercadoria. Como instrumento cultural forte, a música estabelece uma relação de proximidade com seu ouvinte. [Adorno, 1983]. Não se pode negar que a música tem um papel fundamental na cultura, sendo um poderoso instrumento social, cultural e político.

Igualmente o mercado de vídeo teve vital importância para a construção das atuais sociedades. Seu crescimento aconteceu no auge dos anos 80, com a expansão das grandes salas de cinema. A televisão veio em seguida tomando conta das casas e criando novos hábitos sociais e culturais, refletindo o que as montagens cinematográficas exibiam.

O processo de digitalização de conteúdos e o uso intensivo dos formatos digitais com técnicas de compressão de dados formaram, em conjunto, o ambiente propício para o surgimento das primeiras redes P2P. Dentre eles, os formatos com compressão mais populares são respectivamente o MP3 para áudio e o MPEG-4 para vídeo.

### **2.4.1 MP3**

Criado em 1987, o MP3 se tornou uma das maiores revoluções no mundo do áudio. Batizado inicialmente como Eureka-EU 147, o MP3, foi desenvolvido no conceituado Instituto Fraunhofer, na Alemanha. A idéia por trás do formato era facilitar a transmissão de áudio de

um ponto a outro, de forma mais rápida e eficiente do que as existentes até então. Sua aplicação inicial era prevista para equipamentos de telefonia de grande porte. [http\_12]

Alguns meses após o início das pesquisas e os primeiros resultados, a “Moving Pictures Experts Group” (MPEG), convidou o laboratório do instituto alemão para trabalhar em parceria com a “International Standards Organization” (ISO), entidade que responde pelos padrões de qualidade industriais seguidos mundialmente. A parceria seria para o trabalho conjunto em projetos de pesquisa para codificação de áudio e vídeo. [http\_12]

Nesta parceria, foram desenvolvidas diversas técnicas de codificação de áudio e vídeo. Entre as criações conjuntas estava o MPG Audio Layer 3, popularmente conhecido como MP3. Com uma abordagem diferenciada na codificação, o áudio original era analisado por um software especialista que se baseava no princípio de áudio perceptual. Este princípio só considerava as frequências sonoras perceptíveis no espectro da audição humana, sendo descartadas as demais frequências no momento da codificação. Aliados a este princípio, foram adicionados algoritmos complexos de compressão de dados a fim de diminuir ainda mais seu tamanho. Como resultado, um arquivo no formato MP3 além de possuir ótima fidelidade, ocupava dez vezes menos espaço do que seu original, cumprindo o objetivo de seus idealizadores. [http\_12]

Em 1996, os computadores pessoais passavam por uma revolução significativa que contribuiu para adoção do MP3. Foram lançadas famílias de processadores com um alto poder computacional, como o Pentium da Intel. Além do aumento dos discos rígidos e a popularização das redes de banda larga.

Tudo isso marcou o grande salto para expansão do formato MP3. Em 1997, programas codificadores de MP3 surgiram na Internet e tocadores (players) se multiplicavam. Desde então o MP3 se disseminou pelo mundo com uma velocidade avassaladora e se tornou o formato preferido de áudio dos usuários das redes P2P para trocarem arquivos na Internet.

Diversos outros formatos surgiram com baseados nas idéias do MP3, sendo os mais famosos Windows Media Áudio – WMA (Microsoft), Ogg Vorbis - OGG(Xiph.Org Foundation)[Formato Livre], Real Áudio- RA – (Real Network), VQF – (Yamaha).

#### **2.4.2 MPEG-4**

O formato MPEG-4, assim como o MP3 foi subsidiado pela “Moving Pictures Experts Group” (MPEG) juntamente com a ISO, onde foi especificado e posteriormente adotado como padrão. Vários centros de pesquisa no mundo contribuíram para a especificação do MPG-4 tendo suas pesquisas finalizadas em 1998 e tornando-se um padrão internacional em 2000.

[KOENEN, 1998]

Grandes melhorias foram realizadas, se comparados ao seu antecessor o MPG-2 (adotado como padrão de vídeo nos DVDs), como a adição do suporte de gráficos 3D, sistema de proteção de direitos autorais (“Digital Rights Management”), menus iterativos, dentre outras. Uma das mais importantes era seu tamanho comparado ao seu antecessor (MPG-2) podendo o mesmo conteúdo ser até seis vezes menor em MPG-4. Além de ter seu uso inicial previsto em transmissões de TV digital.

Empresas como a Apple e a Microsoft adotaram logo a tecnologia, em meados de 2002, para seus tocadores (players) e criaram formatos proprietários baseados na sua especificação.

Além dos desenvolvedores independentes que utilizaram a especificação do MPEG-4 para acrescentar outras capacidades e criarem subaplicações, expandindo ainda mais seus recursos. Dentre os mais populares, o Divix da Divix Lab, criado em 1999, e o Xvid, formato mantido pela comunidade de software livre da internet. A adoção desses formatos de arquivos de vídeo foi imediata nas redes P2P.

### 3 REREFENCIAL TEÓRICO II

#### 3.1 DESAFIOS DA INDÚSTRIA DO ENTRETENIMENTO

A troca de conteúdos de forma ilegal tem se tornado um dos principais desafios para indústria do entretenimento. No Brasil a *ABPD (Associação Brasileira de Empresas de Disco)*, afirma que a troca sem autorização de arquivos na Internet é hoje uma das principais causas da queda nas vendas de CDs. Somente esta forma ilegal de aquisição de fonogramas ocasionou queda de 21% entre 2000 e 2004[http\_1]. No contexto global, a situação não é diferente; em um recente relatório apresentado em 2005 pela *IFPI (Federação Internacional da Indústria Fonográfica)* demonstrou-se que aproximadamente 900 milhões de arquivos musicais não autorizados estão disponíveis nas redes P2P[http\_2].

Na área de direitos de propriedade intelectual as perdas também já são sentidas. A *RIAA (Associação da Indústria Fonográfica da América)*, que defende os direitos de propriedade intelectual fonográficos, nos EUA, diz também que milhões de dólares são perdidos anualmente pelo tráfico de arquivos de forma não autorizada nas redes P2P [http\_4]. Porém, vale ressaltar que, no Brasil, apesar de não se ter estudos específicos sobre o tráfico não autorizado de conteúdos, a *APDIF (Associação Protetora dos Direitos Intelectuais Fonográficos)* retirou da Internet em 2004 mais de 4113 sites com conteúdo digital sem autorização. [http\_5]

Nas redes P2P trafega-se todo o tipo de conteúdo, inclusive filmes. Somente o prejuízo, estimado, causado pelo tráfico indiscriminado de arquivos cinematográficos chegou a 1 bilhão de dólares por ano, segundo a *MPA Europa (Motion Picture Association)*. [http\_4]

Os usuários dessas redes se multiplicam exponencialmente, segundo dados do Ibope NetRatings. Os softwares mais conhecidos que utilizam protocolos P2P variados como o "eMule", em novembro de 2005 no Brasil, teve mais de 1,2 milhões de usuários. Cerca de 552 mil usuários usaram a rede "Gnutella" e mais de 301 mil o "Kazaa", enquanto "LimeWire" e "Soulseek" registraram 272 mil e 189 mil usuários, respectivamente, para o mesmo período. [http\_3]

As medidas e ações legais para a proibição da troca desses arquivos mostram-se insuficientes. São milhões de usuários de diversas redes com protocolos variados em quase todos os países do globo, trafegando conteúdos distintos simultaneamente [http\_6, http\_7]. A natureza dos protocolos utilizados e as características físicas da própria Internet dificultam ainda mais qualquer tentativa de se controlar o que é trafegado, em que rede P2P e qual usuário está por trás do seu uso. E mesmo com todas as tentativas de controle alguns protocolos já programam técnicas para ocultar rastros. [Wen, 2004]

Para entender melhor como as redes P2P chegaram a este patamar de uso e suas principais características, serão analisados os conceitos essenciais de redes de computadores e dos sistemas distribuídos.

### **3.2 REDE DE COMPUTADORES**

As redes de computadores já fazem parte do dia a dia das pessoas nos escritórios e nos lares; hoje em dia são tão comuns em bancos, instituições de ensino que chegam a ser quase invisíveis. E uma das redes mais conhecidas na atualidade, é a Internet - a rede mundial de computadores. Através dela usuários em todo o mundo compartilham informações de forma rápida e eficiente. [ Abbate J.,2000 ]

Desde a guerra fria, onde os primórdios da Internet foram calçados nas redes militares, já se interconectavam um ou mais dispositivos com regras pautadas em protocolos. [Abbate J. 2000; Tanenbaum, A. 2003]. Este conjunto de regras, os protocolos, já ditavam as normas da comunicação das redes e estabeleciam os eventos de envio, recebimento e localização de informações. [Tanenbaum, A. 2003]

### **3.3 SISTEMAS DISTRIBUÍDOS**

Os nós são entidades computacionais que podem interagir entre si, e cada um representa um computador individual ou aplicações, ligadas entre si, indicando que estas fazem parte de uma mesma rede de informações. Estes dispositivos compartilham conteúdos com regularidade, independente da direção da comunicação ou posição dos elementos. [Minar, N. 2001]

Sistema Distribuído é todo sistema que pode dividir suas tarefas de forma transparente, independente do número de dispositivos (nós) ligados ao sistema colaborativo através de uma rede de comunicação. Toda tarefa pode ser dividida entre esses nós idênticos em funcionalidade, que se comunicam através de um protocolo. [Coulouris, G. 2005; Tanenbaum, A. 2003]

Os Sistemas Distribuídos possuem três características chave: tolerância à falha, escalabilidade e baixo custo de manutenção, como visto na tabela 3.

<b>Características Chave</b>	<b>Descrição</b>
<b>Tolerância à Falhas</b>	Um ou mais nós podem falhar ( hardware, software) sem comprometer o sistema.
<b>Escalabilidade</b>	A qualquer momento podem sair e entrar nós crescendo ou diminuindo essa rede sem complicações para o sistema.
<b>Baixo Custo de Manutenção</b>	O custo computacional / infra-estrutura é compartilhado entre os nós. O sistema aproveita a conectividade e o hardware existente do nó integrante.

**Tabela 3** – *Sistemas Distribuídos – Características* [ Coulouris, G. 2005; Kindberg,T .2005]

Dentro dos sistemas distribuídos uma caracterização importante é a descrição das ligações realizadas. As topologias, como são chamadas, representam os caminhos e as direções de interligação dos nós.

Cada topologia tem um objetivo distinto, podendo ser desenhada para obter velocidade, eficiência, pertinência, etc. [Minar, N. 2001].

### **3.3.1 TOPOLOGIAS DE SISTEMAS DISTRIBUÍDOS**

Nas topologias vários aspectos são analisados como o acesso físico, organização, fluxo de informações e etc. As classificações das topologias se diferem pela forma que os nós estão correlacionados. Para o presente estudo, foram divididas duas categorias topológicas: as puras e

as híbridas, levando-se em conta sempre o fluxo da informação para definição de sua classificação. [Minar, N. 2001].

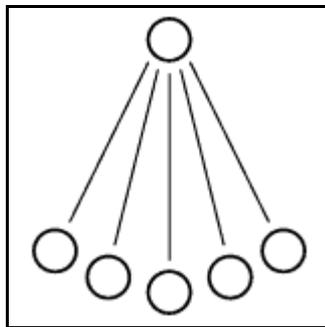
São definidas como topologias puras as que não mesclam outras topologias na sua composição. Já as híbridas são topologias que são combinadas com outras para oferecer um sistema com características mistas.

Através destas misturas, por exemplo, é possível criar uma topologia híbrida que utilize duas topologias puras [Minar, N. 2001].

### **3.3.2 TOPOLOGIAS DE SISTEMA SISTEMAS DISTRIBUÍDOS PURAS**

#### **3.3.2.1 TOPOLOGIA CENTRALIZADA**

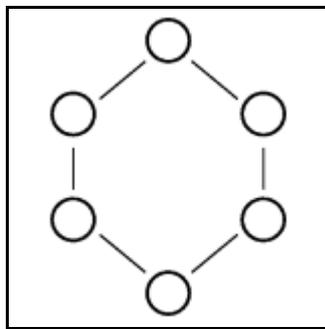
A Topologia Centralizada possui padrões visíveis da arquitetura cliente/servidor, onde todas as informações são agrupadas em um único servidor, que pode possuir muitos usuários conectados diretamente, recebendo e enviando informações simultaneamente [Minar, N. 2001].



**Figura 12 - Topologia Centralizada** (Minar, N. 2001)

### 3.3.2.2 TOPOLOGIA EM ANEL

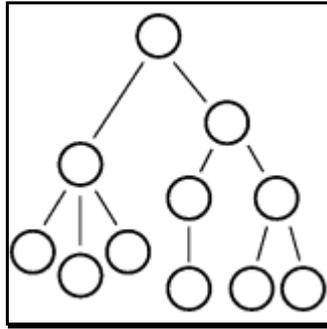
Na topologia, centralizada existe um único dispositivo centralizado; esse dispositivo, por ser centralizado, não suporta a carga elevada de todas as requisições. Uma solução comum para esse problema é utilizar um conjunto de máquinas arrumadas em forma de círculo (anel), objetivando balancear essa demanda. Todos os nós passam a ter funcionalidades idênticas e um protocolo coordena o fluxo, gerenciando e unindo as informações. Desta maneira, a rede estará preparada para suportar as falhas de um ou mais nós (Minar, N. 2001).



**Figura 13** - Topologia em Anel (Minar, N. 2001)

### 3.3.2.3 TOPOLOGIA HIERÁRQUICA

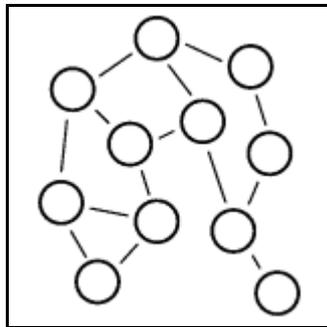
Trata-se da topologia peculiar dos sistemas distribuídos, normalmente utilizada com protocolos simétricos que propagam a informação em cascata através da árvore de nós. Os nós-pais agem como pequenos servidores propagando através de sua espinha dorsal(*backbone*) informação para os seus nós-filhos.(Minar, N. 2001).



**Figure 14 - Topologia Hierárquica (Minar, N. 2001).**

#### **3.3.2.4 TOLOGIA DESCENTRALIZADA**

Topologia em que os nós se comunicam simetricamente em todas as direções. Todos os nós são autônomos e idênticos em funcionalidade, possuindo alto índice de tolerância à falhas. Esta topologia, por exemplo, é utilizada no roteamento de informações da própria Internet (Minar, N. 2001).

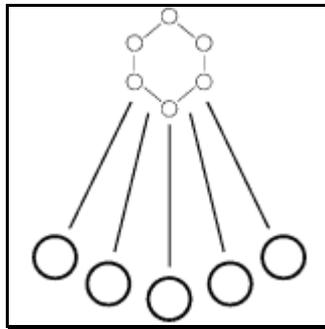


**Figura 15 - Topologia Descentralizada (Minar, N. 2001)**

### 3.3.3 TOPOLOGIA DE SISTEMAS DISTRIBUÍDOS HÍBRIDAS

#### 3.3.3.1 TOLOGIA CENTRALIZADA + TOPOLOGIA ANEL

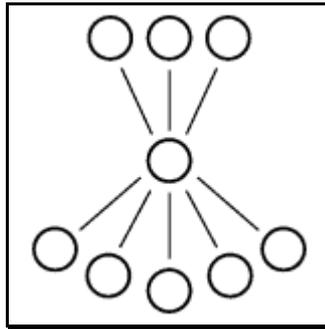
Sabendo-se que as topologias em anel são comumente utilizadas para balancear a carga e criar um sistema que suporta falhas, é possível mesclar-se à Topologia Centralizada para prover mais benefícios (Minar, N. 2001). O ganho é observado com a simplicidade de controle da Topologia Centralizada, aliada à robustez da Topologia em Anel no centro da rede (Minar, N. 2001).



**Figura 16** - *Topologia Centralizada + Topologia Anel* (Minar, N. 2001).

#### 3.3.3.2 TOLOGIA CENTRALIZADA + TOPOLOGIA CENTRALIZADA

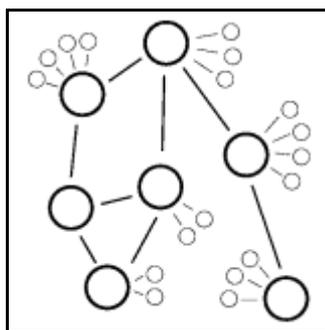
Esta topologia pode ser caracterizada com a união de várias topologias centralizadas, onde mais de um servidor está na rede compartilhando o direito de responder às requisições dos clientes. Esta estrutura é muito comum em servidores web, que são acessados por vários usuários. Para eles, o uso de mais de um servidor é transparente e a aplicação principal (protocolo) do sistema coordena por qual caminho cada cliente irá acessar e qual servidor se conectar (Minar, N. 2001).



**Figura 17** - *Topologia Centralizada + Topologia Centralizada (Minar, N. 2001)*

### 3.3.3.3 TOLOGIA CENTRALIZADA + TOPOLOGIA DESCENTRALIZADA

A mais promissora das topologias, com uma arquitetura avançada de topologias centralizadas dentro de uma topologia descentralizada (*Minar, N. 2001*). Um exemplo simplista é o serviço de e-mail, que possui servidores centralizados que respondem aos endereços nele registrados. Por sua vez, eles enviam mensagens a outros numa topologia descentralizada através do servidor de e-mail de destino. Conseqüentemente, recebe o e-mail e encaminha a um ou mais clientes (*Minar, N. 2001*).



**Figura 18** - *Topologia Centralizada + Topologia Descentralizada (Minar, N. 2001)*

### 3.4 REDES P2P

#### 3.4.1 DEFINIÇÕES DE REDES P2P

Entendendo agora as principais topologias dos sistemas distribuídos, pode-se definir então o que são as redes P2P. Apesar de muitas definições já terem sido formuladas sobre as redes P2P, para melhor o entendimento de suas características, optou-se neste trabalho por tabular as principais definições encontradas na literatura e seus respectivos autores.

Definições	Autores
Ausência de controle centralizado	Stoica e Balakrishnan (2001)
Nós equivalentes em funcionalidades	
Auto-organização	Rowstron e Druschel (2001)
Adaptabilidade	
Escalabilidade	
Comunicação Simétrica	
Participação voluntária dos nós	Saroiu e Gribble (2002)
Contribuição com recursos de infra-estrutura	

**Tabela 4 - Definições P2P / Autores**

Uma das características mais relevantes das redes P2P é a ausência do controle centralizado, como demonstra o primeiro item da tabela 2 de definições P2P. Essa afirmação mostra-se verdadeira mesmo quando sua categoria topológica é pura, utilizando-se de servidores centrais para a indexação de conteúdos. [http\_8]. Convém também observar que todos os nós cliente são equivalentes em funcionalidade (*Stoica e Balakrishnan, 2001*).

As características de auto-organização e adaptabilidade de estrutura surgem quando um novo nó une-se à rede P2P em funcionamento e passa a participar da comunicação e do roteamento de informações através do protocolo (*Rowstron e Druschel, 2002*).

A capacidade simétrica que as redes P2P têm na sua comunicação faz com que um nó sempre procure o menor caminho (melhor roteamento), equalizando sempre as distâncias entre eles a fim de proporcionar um melhor desempenho (*Rowstron e Druschel, 2002*).

Unidas a essas características, a escalabilidade ainda permite o crescimento ou a redução da rede sem interrupções da comunicação. Isso ocorre caso um nó da rede P2P esteja comunicando uma fração de um determinado conteúdo e falhar por algum motivo. Imediatamente, o protocolo procurará um nó ativo na rede para suprir essa ausência que contenha esta fração (*Rowstron e Druschel, 2002*).

E a última característica é que todos os nós participam da rede P2P voluntariamente (sem troca de valor monetário), seja compartilhando conteúdos ou somente recebendo informações de acordo com as regras ditadas no protocolo. Além da cessão de poder computacional (processamento) e sua largura de banda para uso de toda a rede P2P. (*Saroiu e Gribble, 2002*)

Reunindo as idéias em torno da definição das redes P2P, pode-se dizer que são sistemas distribuídos, construídos em cima de uma rede de computadores já estabelecida. As características físicas da rede (hardware), conectividade e localização geográfica não influenciam isoladamente. O funcionamento da rede P2P não é dependente de um único dispositivo central ou aplicação matriz, possuindo na sua sistemática, um alto grau de tolerância à falhas. (*Saroiu e Gribble, 2002; Rowstron e Druschel, 2002; Stoica e Balakrishnan, 2001*)

### **3.4.2 FUNCIONAMENTO DAS REDES P2P**

Uma rede P2P normalmente objetiva que conteúdos sejam trafegados pelo maior número de usuários possíveis. Para isso, o processo de aquisição se inicia com a busca de um determinado conteúdo, que tem de ser encontrado de alguma maneira e em seguida referenciado. [STEFAN , 2002]

Após o seu referenciamento, que dirá onde o conteúdo está (IP ou rota para encontrar o nó), se inicia então o processo de transferência. Este processo geralmente acontece com a divisão do conteúdo em partes menores, dividindo o trabalho de envio com todos os nós que tenham o conteúdo em comum.[STEFAN , 2002]

### **3.4.3 NOMECLATURAS DAS REDES P2P**

Na literatura (*Ge and Towsley, 2003; Schollmeier 2001*), foram encontradas algumas nomenclaturas para dividir as redes P2P. No presente estudo, foram analisadas as do autor Schollmier, por dividir de forma topológica a classificação das redes P2P com base em três propriedades: Servidor, Método de Busca dos conteúdos, e de que forma sua Infra-estrutura lógica é composta.

Para melhor entendimento foram tabuladas para este estudo as propriedades de Schollmeier :

	<b>P2P Pura</b>	<b>P2P Híbrida</b>
<b>Servidor</b>	Centralizado	Não Utiliza
<b>Busca</b>	Possuem listas de conteúdos	Possuem parte da lista de conteúdos
<b>Infra-Estrutura</b>	Semi-Centralizada	Distribuída

**Tabela 5** - *Nomenclaturas de Schollmeie*

#### **3.4.4 REDE P2P PURA**

Uma rede P2P é considerada pura quando se utiliza de servidores centrais para conectar o nó-cliente (software) a outros nós. Neste caso, os servidores funcionam apenas para armazenar os endereços dos usuários on-line e guardar temporariamente a lista dos conteúdos disponibilizados por eles. Uma vez que o conteúdo de interesse tenha sido listado pelo servidor e selecionado pelo cliente (software), o servidor já não se faz necessário, pois o tráfego do conteúdo - roteamento e operações de comunicação - é realizado diretamente ponto a ponto, ou seja, de cliente para cliente (Schollmeier,2001).

#### **3.4.5 REDE P2P HÍBRIDA**

Com uma estrutura distribuída não é necessário o uso de servidores centrais para seu funcionamento. Todos os componentes da rede como a lista de arquivos e as ligações da infraestrutura lógica são divididas entre todos os nós ativos (Schollmeier,2001).

Na prática, as redes P2P têm eventos similares, como o ato de procurar, referenciar e adquirir conteúdos. Existe variação apenas da implantação de cada protocolo, que pode estruturar sua topologia de forma mista, a fim de manipular o tráfego de informações entre os nós para obtenção de ganhos (*Minar, N. 2001*).

### **3.5 PROTOCOLOS P2P**

Todos os protocolos P2P procuram diferenciar-se em algum quesito como velocidade, número de conteúdos disponíveis, sigilo de identidade e etc. Neste estudo, foram analisados os principais protocolos ativos e suas características.

O entendimento dos protocolos e suas características se fazem necessário, pois o protótipo produzido neste trabalho utilizou-se das suas especificações para seu desenvolvimento. Não foram considerados softwares-clientes, já que existem dezenas deles para cada protocolo e alguns deles ainda suportam mais de um protocolo.

#### **3.5.1 PROTOCOLO FASTTRACK**

O protocolo *Fasttrack* Surgiu durante o fim da primeira geração de redes P2P, produzido por vários programadores incluindo um membro do time que produziu o Skype. Foi introduzido em março de 2001 pela empresa Alemã (*Consumer Empowermet*) e hoje sob a licença (*GPL*).

[http\_16]

Sua principal característica é de uma implementação especial chamada "supernodes", uma nomenclatura dada aos nós especiais que aumenta consideravelmente a escalabilidade da rede, cedendo mais banda e armazenamento. Esta implementação distribui a carga de indexação dos

conteúdos entre os nós de forma inteligente. Os nós que têm mais conectividade (largura de banda) indexam os arquivos funcionando como "servidores" temporários. Isso ajuda aos nós lentos a não ficarem sobrecarregados, “segurando” a rede, aumentando assim, como um todo, a velocidade da rede. [http\_16]

Estima-se que em 2006 a rede tenha aproximadamente 2.7 milhões de usuários simultâneos por segundo utilizando este protocolo. [http\_16]

### **3.5.2 PROTOCOLO EDONKEY**

A rede *Edonkey*(também chamada de rede *EDonkey2000* ou *ed2k*) foi criada pela empresa Alemã MetaMachine no final de 2000, incentivada pelo sucesso do cliente Napster. Tinha como desafio o compartilhamento de grandes arquivos, da ordem de gigabytes. [http\_17]

Sendo uma rede descentralizada, os conteúdos não são guardados no servidor central e sim distribuídos diretamente entre os usuários com base nos princípios peer-to-peer. Utiliza-se de servidores para indexar os arquivos disponíveis nos nós conectados em cada servidor. Uma das suas principais características é a capacidade de distribuição de grandes arquivos, superiores a 9.8 megabytes. O protocolo utiliza técnicas de segmentação inteligente do conteúdo em partes menores, facilitando assim a velocidade de downloads nos nós clientes. [http\_17]

No meio de 2005 a rede tinha mais de 3 milhões de usuários simultâneos por segundo, compartilhando quase 2 bilhões de arquivos com a ajuda de pouco mais de 200 servidores para indexar os arquivos. [http\_17]

### 3.5.3 PROTOCOLO GNUTELLA

O protocolo *Gnutella* foi desenvolvido por Justin Frankel e Tom Pepper da empresa Nullsoft no início de 2000, logo após a empresa ser adquirida pelo AOL. Em 14 de Março de 2000, o primeiro programa estava disponível para download. A AOL tirou do ar devido a problemas legais, mais era tarde, pois a própria comunidade de programadores da Internet fez engenharia reversa no software, reescrevendo assim o protocolo e liberando com uma licença de uso livre (open-source). [http\_18]

A grande diferença técnica do protocolo *Gnutella* é que as buscas funcionam baseadas em uma topologia anel. O cliente que tem implementado o protocolo lança a pesquisa para um nó que, por conseguinte envia a outro e assim sucessivamente. Dessa forma, o protocolo do *Gnutella* capacidade de trafegar arquivos grandes com a técnica de partilhamento. [http\_18]

Estima-se que em 2006 a rede tenha aproximadamente 2.2 milhões de usuários simultâneos por segundo utilizando este protocolo [http\_18].

### 3.5.4 PROTOCOLO BITTORRENT

Esse protocolo foi criado em *Bram Cohen* em 2003, e tem se popularizado a cada dia. *Bittorrent* é um protocolo que permite receber arquivos que estão indexados em sites da Internet, não necessitando de servidores para nenhuma tarefa da comunicação. O protocolo do Bittorrent originalmente foi escrito em Python, uma linguagem de programação recente e poderosa. [http\_19]

Esse protocolo introduziu o conceito de cotas, qual o usuário tem direito de receber a mesma quantidade de bytes que envia. Esta técnica possibilitou downloads de conteúdos mais rápidos. Sendo assim, quanto mais pontos (nós) tentem baixar um determinado conteúdo, mais rápida será sua distribuição. [http\_19]

Estima-se que mais de 3 milhões de usuários simultâneos por segundo estejam utilizando este protocolo. [http\_19]

### **3.5.5 PROTOCOLO FREENET**

O protocolo *FreeNet* foi desenvolvido sobre licença livre. O seu autor *Ian Clarke* concebeu sua idéia em julho de 1999 quando ainda era estudante na universidade de Edimburgo, publicando no artigo “Sistema Descentralizado Distribuído de Armazenamento e de Recuperação de Informação”. Logo após a publicação do artigo, *Clarke* e um número pequeno de voluntários começaram a implementá-lo, dando origem assim ao que mais tarde se tornaria a rede *Freenet*. [http\_20]

Mais tarde a rede *Freenet* se tornaria a rede P2P com a capacidade peculiar de deixar os seus usuários anônimos usando técnicas e algoritmos complexos de repasse de partes de conteúdos por vários nós. Essas técnicas dificultam ao extremo qualquer possibilidade de identificação de seus usuários. A *Freenet* então se tornou a primeira “Darknet”. [http\_20]

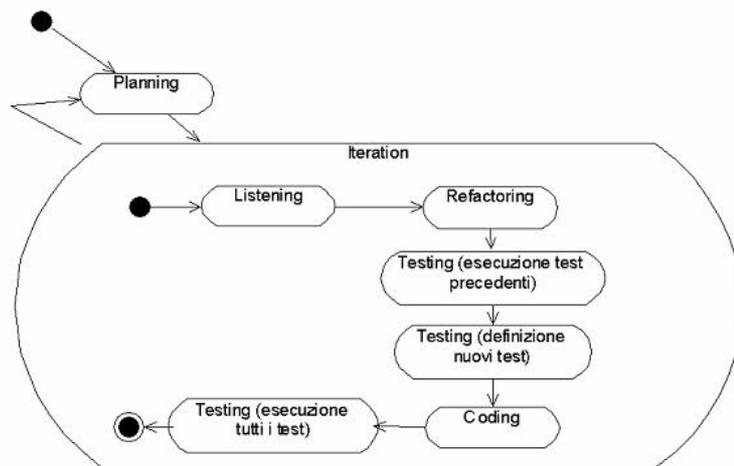
## **4 METODOLOGIA**

### **4.1 MOTIVAÇÃO**

A motivação do protótipo é mensurar quanto conteúdo é circulado dentro de uma gama de servidores ou nós de uma rede P2P.

### **4.2 METODOLOGIA DE DESENVOLVIMENTO DO PROTÓTIPO**

Para este protótipo foi escolhida a metodologia de *Extreme Programming* ou XP. Esta metodologia descreve um processo minimalista onde existe pouca burocracia no desenvolvimento. Normalmente utilizada em equipes pequenas, de até 10 desenvolvedores, trabalhando em interações curtas, produzindo software incrementalmente. E a análise dos requisitos é feita à medida em que uma nova necessidade surge. Na figura 19 é exibido o fluxo de iterações na metodologia XP.



**Figura 19 - Fluxo metodologia XP**

Esta metodologia se apóia em um contínuo refinamento do projeto e da implantação do código aplicado a certos princípios. No caso específico da implantação do protótipo deste estudo, algumas premissas não foram utilizadas, pois só existia um programador no projeto.

Premissas de XP utilizadas no protótipo:

- **User Stories:** requisitos são elaborados na forma de cenários de operação do software pelo usuário, e o software é implantado a partir destes cenários.
- **Refactoring:** o código é continuamente refeito para que se reduza à sua forma mais simples e legível possível.
- **Testes de Unidade:** a programação é iniciada por testes de unidade, que são implantados antes do código que devem testar. A coleção de testes permite aos desenvolvedores alcançarem garantias o suficiente para continuamente implantando código, de forma que o código possua poucos erros.
- **Documentação:** a documentação do sistema deve ser mantida direto no código fonte, e deve ser curta e direta, forçando o desenvolvedor a escrever um código legível sem complexidades.

### 4.3 REQUISITOS DO PROTÓTIPO

Os requisitos do protótipo estão diretamente ligados a sua própria capacidade de cumprir as metas proposta na motivação. A designação dos requisitos foi focada nos requisitos funcionais críticos. Consideram-se requisitos críticos para este protótipo todos aqueles que numa fase avançada do projeto significassem alto grau de modificação do código em produção, caso não fossem planejados anteriormente.

Para este protótipo foram considerados os seguintes requisitos críticos:

- **Desempenho:** O protótipo deve conseguir interagir pelo menos com 10 servidores simultâneos a fim de criar uma mostra estatística mínima para o objetivo fim. E suportar a execução de duas ou mais estâncias do protótipo durante o período de coleta de dados.
- **Portabilidade:** O protótipo deve ser portátil para qualquer para as plataformas X86, X86-64 e RISC; e qualquer sistema operacional utilizado em larga escala como o Windows, Unix e Linux.
- **Internacionalização:** Devido à diferença da língua portuguesa e suas peculiaridades (acentos), é necessário um suporte a formatos de caracteres latinos como o ISO 8859-1 (Latin-1)[http1].
- **Software Livre:** Desenvolver o protótipo com uma licença livre, que permita a sua redistribuição de igual forma. Além de ser desenvolvido numa linguagem que possua um compilador ou máquina virtual de igual licença.

#### **4.4 CONCEPÇÃO DO PROTÓTIPO**

Durante a concepção do projeto foi necessário optar por qual protocolo P2P se utilizaria. Para fins de estatística e por possuir uma estrutura mais centralizada, optou-se pelo uso do protocolo Edonkey. Além do mesmo ser largamente utilizado pelo público brasileiro[http], sua estrutura de busca de arquivos e seu modelo de comunicação tornou-se pertinente para aplicação deste protótipo.

#### **4.5 PROJETOS ANTERIORES**

Após larga pesquisa pela Internet, procurou-se por projetos que objetivassem a coleta de dados estatísticos das redes que utilizassem o protocolo do Edonkey. Não foi encontrada nenhuma ferramenta ou biblioteca que auxiliasse diretamente nesta tarefa. Porém foi encontrada a empresa americana BigChampagne[http\_21], que alega possuir uma ferramenta com este objetivo, mas com registro de patente ainda pendente.

A BigChampagne atualmente tem grande carteira de clientes incluindo a revista musical americana Billboard, sites como o iTunes, MSN Music, Yahoo, além das grandes gravadoras de todo o mundo. Segundo a própria empresa, seu software estatístico varre dezenas de redes P2P em todo o mundo ao mesmo tempo, com diversos servidores conectados para coletar dados sobre conteúdos multimídias trocados ou pesquisados. A figura 20 apresenta a tela do site da empresa.

**new here?**  
**60SECOND Champagne**

THE STANDARD MEASURE OF THE WORLD'S MOST POPULAR DIGITAL MEDIA...  
...SINCE THE LATE 20TH CENTURY

"BigChampagne data is the most widely accepted."  
- Rolling Stone

In the 21st century, consumers are acquiring more music online than on CD. Every month, tens of millions of people in the U.S. alone turn to the Net for new music. These are the week's ten most popular songs...

**topSwaps**  
.: Top10 (all) .: April 17, 2006 to April 24, 2006

Wk	LW	TW	Artist	Track
#02	#01	↑	Rihanna	SOS
#03	#02	↑	Ne-Yo	So Sick
#07	#03	↑	Ne-Yo	When You're Mad
#09	#04	↑	Pink	Stupid Girls
#04	#05	↓	Chris Brown	Yo
#05	#06	↓	Sean Paul	Temperature
#08	#07	↑	Pussycat Dolls	Beep
#15	#08	↑	Eminem	Shake That
#06	#09	↓	Bubba Sparxxx	Ms New Booty
#10	#10	↓	Busta Rhymes	Touch It

>> [click here to see more detail](#)

© 2001-2005 BigChampagne, LLC - all rights reserved.

Figura 20 - Tela do Site BigChampagne

As estatísticas são cobradas por unidades pesquisadas e a ferramenta não tem qualquer especificação de domínio público para seu comparativo com este protótipo. A única informação disponível para o público é o conjunto de músicas mais procuradas na Internet, separadas por categorias.

## 4.6 O PROTÓTIPO

### 4.6.1 LINGUAGEM DE PROGRAMAÇÃO

Para o desenvolvimento deste protótipo era necessária a de uma linguagem de programação e para isso, foram analisados alguns critérios:

- Biblioteca que auxiliasse na escrita de protocolos de rede assíncronos;
- Suportasse um desenvolvimento ágil;
- Sintaxe flexível e minimalista;
- Suporte a orientação a objetos;

- Fácil aprendizado.

Além dos critérios adotados acima, procurou-se optar pela linguagem de programação que fosse mais adequada para a escrita do protocolo do Edonkey adaptado. Dentre as dezenas de linguagens disponíveis no mercado, optou-se pela linguagem Python.

#### **4.6.2 LINGUAGEM PYTHON**

Criada pelo holandês Guido *van Rossum*, apresentou diversas características interessantes para este protótipo. Possuía tipos dinâmicos, gerenciamento de memória automático e suporte a diversos paradigmas de programação como orientação a objetos, programação estruturada e programação funcional. Além de sintaxe flexível e minimalista; tendo excelente portabilidade e de fácil aprendizado a tornaram uma excelente opção para o desenvolvimento deste protótipo. [http\_22]

Outro ponto que justificou sua escolha foi o fato de possuir uma excelente biblioteca para a criação de protocolos chamada Twisted. A figura 21 mostra um trecho do código do protótipo escrito na linguagem python.

```

class structStream(structSpecial):
    def byte(self):
        return struct.unpack('<B',self.slice(1))[0]
    def word(self):
        return struct.unpack('<H',self.slice(2))[0]

    def dword(self):
        return struct.unpack('<I',self.slice(4))[0]

    def qword(self):
        return struct.unpack('<II',self.slice(8))[0]

class emuleRead(structStream):
    def __init__(self,data):
        self.data = data
        self.pos = 0
        self.opcodes = [ ('OP_HELLO',0x01),
                          ('OP_HELLOASW',0x4C),
                          ('OP_SERVERMESSAGE',0x38),
                          ('OP_IDCHANGE',0x40),
                          ('OP_SERVERSTATUS',0x34),
                          ('OP_SERVERLIST',0x32),
                          ('OP_SERVERIDENT',0x41),
                          ('OP_SERVERRESULT',0x16),
                          ('OP_FOUNDSOURCES',0x42),
                          ('OP_CALLBACKREQUESTED',0x35),
                          ('OP_CALLBACK_FAILED',0x36),
                          ('OP_SEARCHRESULT',0x33),
                          ('OP_REJECT',0x36),
                          ('OP_UDPSEARCHRESULT',0x99),
                          ('OP_ANSWERSOURCES',0x82),
                          ('OP_FILEREQANSNOFIL',0x48),
                          ('OP_MULTIPACKETANSWER',0x92),
                          ('OP_GLOBFOUNDSOURCES',0x9B)
                        ]

        self.streams = []

    def splitStreams(self):
        EDONKEY_PROTOCOL = 0xE3
        EMULE_PROTOCOL = 0xC5
        SERVER_MESSAGE = 0x38

```

**Figura 21** - Exemplo do Código do Protótipo em Python

Outra boa indicação de que o Python seria uma boa opção para o protótipo, era seu histórico de sucesso neste tipo de aplicação. O moderno protocolo P2P descentralizado do Bittorrent, criado por Bram Cohen, teve suas primeiras versões nesta linguagem.[http\_19]

#### 4.6.3 BIBLIOTECA TWISTED

Uma vez feita a escolha da linguagem Python para a escrita do protocolo do Edonkey, fez-se necessário a verificação de uma biblioteca que comportasse algumas características para tornar ágil a sua implementação como:

- Estrutura assíncrona
- Suporte a listen (escuta) e envio (connect) a múltiplas portas e host

- Boa documentação

Pesquisando na Internet e em listas especializadas de Python, muitos desenvolvedores indicaram a biblioteca Twisted para este fim, afinal até a NASA utilizava a biblioteca para algumas de suas aplicações. [http\_23]

Esta biblioteca foi desenvolvida pelo laboratório da Edgewall Software[http\_24] e com a contribuição de diversos usuários espalhados pelo mundo. Ela foi desenvolvida inteiramente na linguagem python para uso em programas que tivessem que explorar a camada de rede. Esta biblioteca tem uma estrutura de framework, e já suportava diversos protocolos como o HTTP, NNTP, IMAP, SSH, IRC, FTP, etc. Considerando as informações a respeito desta biblioteca, sua adoção foi considerada de imediato, tornando-se então o núcleo para o desenvolvimento do protótipo.

#### **4.6.4 REFERÊNCIAS DO PROTOCOLO EDONKEY**

O criador do protocolo, o alemão Jed McCaleb, não criou uma referência muito didática sobre os eventos de comunicação de seu protocolo. Então procurou-se na Internet outras referências que fossem mais explícitas a esse respeito. Graças a sua popularidade, o protocolo foi muito aplicado em dezenas de programas clientes como o *Emule*. A figura 11 mostra a tela do Cliente Emule que faz uso do protocolo Edonkey.

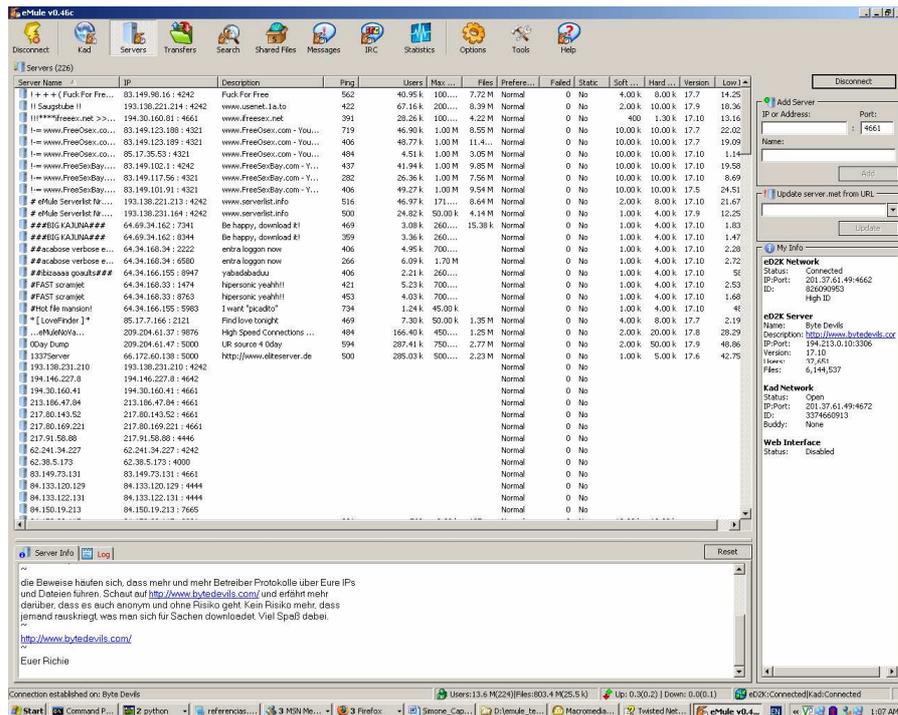


Figura 22 - Tela do Cliente Emule

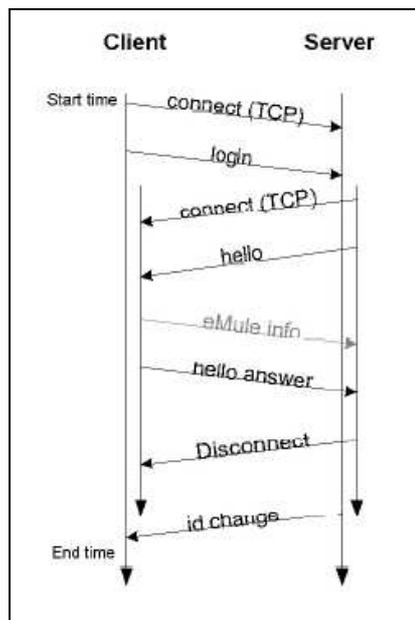
Neste protótipo utilizou-se como referência a especificação do protocolo criado pelos alunos *Yoram Kulbak* e *Danny Bickson*, supervisionados pelo *Prof. Scott Kirkpatrick*, do laboratório *DANSS (Distributed Algorithms, Networking and Secure Systems)* - Universidade de Ciências da Computação e Engenharia da Hebrew University de Jerusalém. Nesta referência o protocolo do *Edonkey* é abordado em suas principais características com pequenas adições específicas do para o cliente Emule[http\_25].

Outras referências foram utilizadas para a escrita do protocolo foi a leitura do código fonte do cliente Emule versão 0.47a implementada na linguagem C++ [http\_26].

#### 4.6.5 ESTRUTURA DO PROTOCOLO

A estrutura do protocolo, como a maioria das redes P2P, utiliza a arquitetura cliente/servidor sendo que cada nó da rede pode se comportar como cliente quando estiver recebendo, e servidor quando estiver servindo informações ou arquivos.

Na figura abaixo é demonstrado o evento de estabelecimento de conexão a um servidor. Como se pode observar o programa cliente, através do protocolo ele envia mensagens para o destino (servidor). E este por sua vez responde, de acordo com o evento iniciado pelo cliente, com uma ou mais mensagens como visto na figura 23.



**Figure 23** - Evento de *Login* com o servidor

## 4.6.6 CRIAÇÃO DE AMBIENTE DE TESTES

Para começar a codificar foi necessária a criação de um ambiente de testes favorável para depurar os possíveis erros durante o desenvolvimento do protocolo na linguagem Python. Isso foi possível graças a utilização do programa Ethereal[http\_27] que dispõem de ferramentas bastante úteis para filtrar os pacotes da placa de rede e depurar as suas mensagens para que elas possam trafegar (vide figura 24).

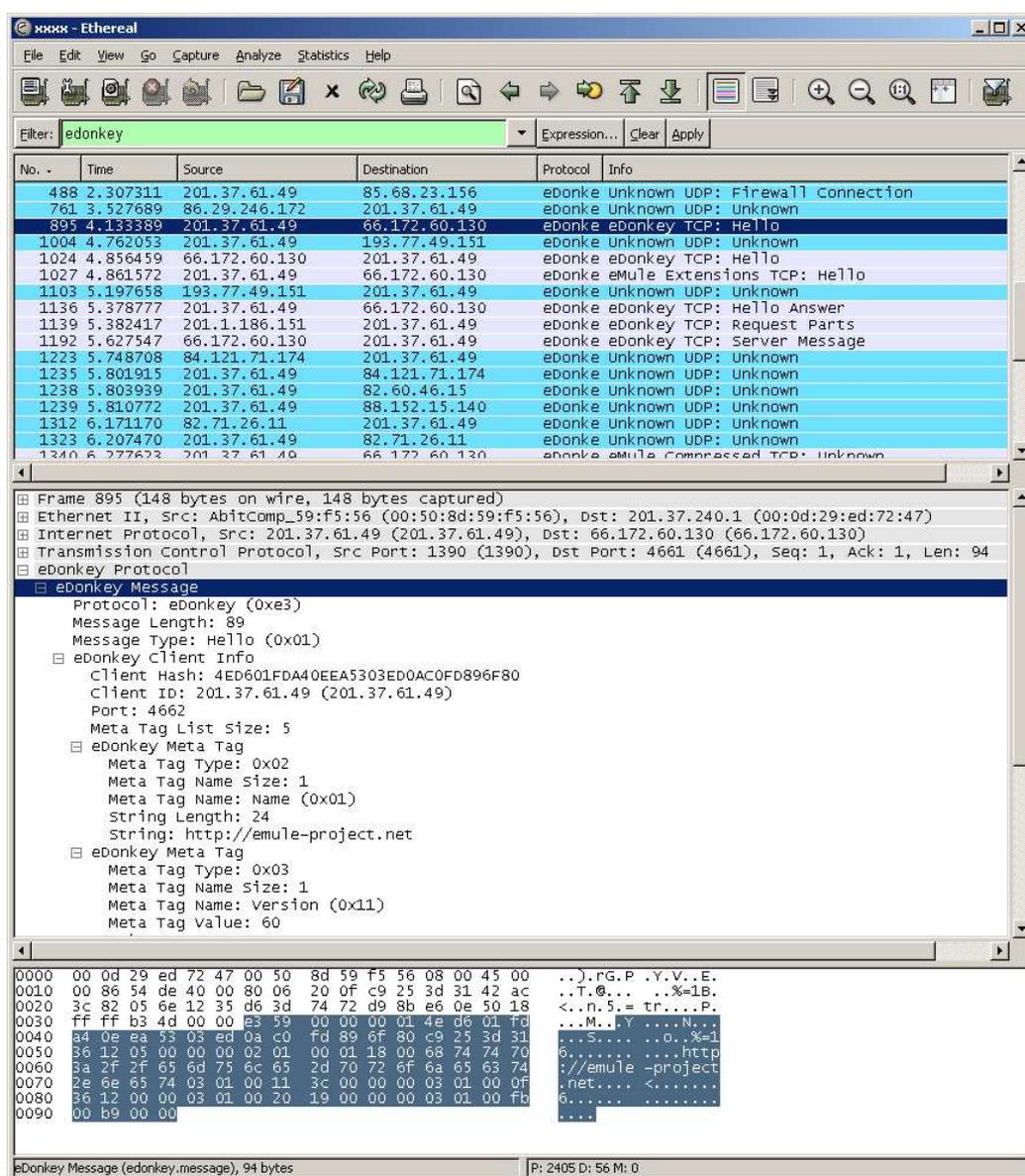


Figura 24 - Tela do software Ethereal

#### **4.6.7 PORTAS E CONEXÕES**

Uma das características complexas no desenvolvimento de protocolos P2P é que por natureza eles são assíncronos, ou seja, não existe fluxo sequencial de informações a seguir. A melhor abstração para um protocolo assíncrono seria a de um software de chat que nunca sabe quantos usuários possui naquele instante ou quando seus usuários enviaram mensagens e para quem. Cabe ao programa estar preparado para “ouvir” as requisições sob demanda, trata-la e executar corretamente as instruções de acordo com cada caso.

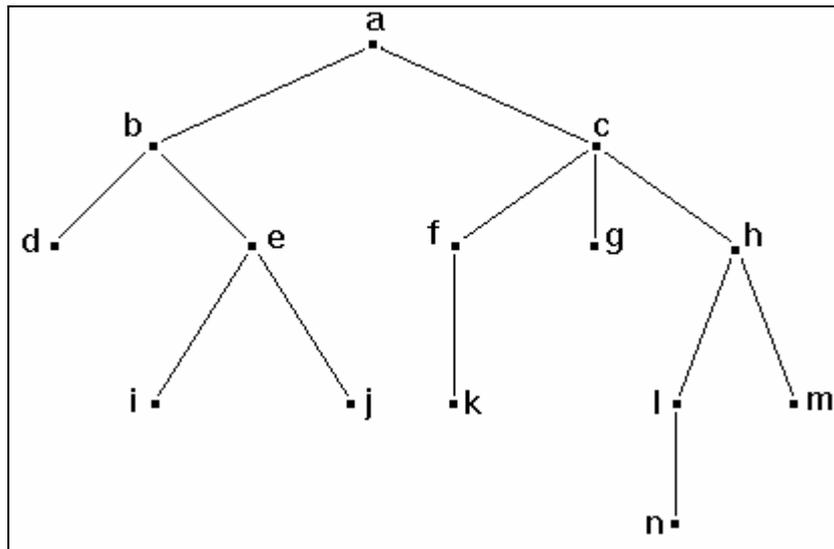
#### **4.6.8 GERAÇÃO DE ESTATÍSTICAS**

As estatísticas para esse protótipo para fins de tabulação foram gerados em XML (eXtensible Markup Language)[[http1](#)] é uma recomendação da W3C(World Wide Web Consortium)[[http2](#)] para gerar linguagens de marcação para necessidades especiais, como o caso desta pesquisa. XML é um subtipo de SGML (Standard Generalized Markup Language) capaz de descrever diversos tipos de dados. Seu propósito principal é a facilidade de compartilhamento de informações.

Tendo a saída das estatísticas em XML é possível processá-las em qualquer software de estatístico com uma adaptação para o seu formato nativo.

#### **4.6.9 BUSCA EM PROFUNDIDADE**

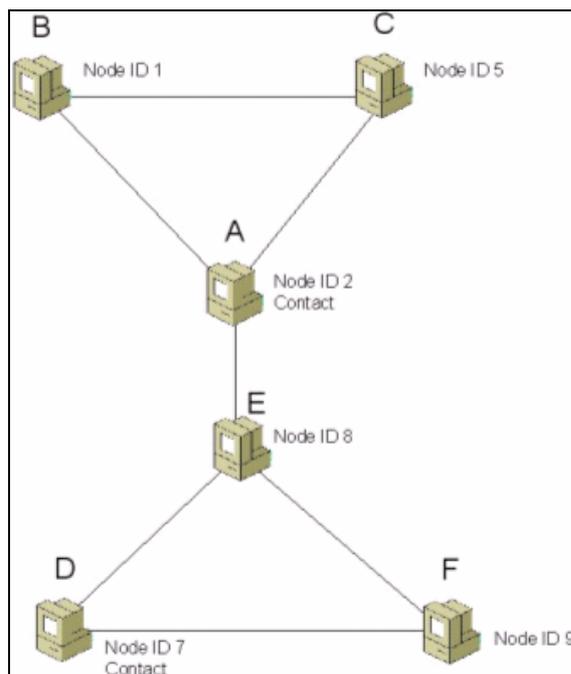
Para a geração das estatísticas se utilizou de uma conhecida disciplina matemática chamada de teoria de grafos. Onde um grafo normalmente é definido sendo como uma coleção de pontos (chamados vértices) conectados em pares por linhas (chamadas de arestas). Como mostrado na figura 25.



**Figura 25** - Representação de um Grafo

A busca em profundidade se caracteriza pela técnica de avançar em níveis da árvore que compõem o grafo. No protótipo em questão foi necessário utilizar-se destas técnicas para poder varrer a rede P2P na sua maior amplitude possível.

A figura 26 mostra a complexidade de um pequeno trecho de uma rede P2P. Onde o nó “A” possui na sua lista de nós que contém o arquivo em circulação os nós “B”, “C”, “E”, “D” e “F”. Por sua vez quando o nós “E” for requisitado a dizer quem são os nós que compõem sua lista, o nó “A” e todos os demais nós virão como retorno. Cabendo ao protótipo identificar que são os nós já existentes para não adiciona-los novamente.



**Figure 26** - Relacionamento dos Nós de uma rede P2P

#### 4.7 CRONOGRAMA DO PROTÓTIPO

O foi produzido do dia 15 de Dezembro até o dia 10 de Junho com um consumo total de 412 horas entre estudo, programação e testes. Como mostra a tabela 6.

<b>Descrição</b>	<b>Horas</b>
Leitura da Documentação do Protocolo	15
Estudo de Código Fonte de Outras Implementações	4
Instalação de Ferramentas	2
Estudo da Linguagem Python	36
Estudo do Pactote Twisted	10
Núcleo ( CORE )	90
Módulo de Escrita de Pacotes	55
Módulo de Leitura da Pacotes	85
Busca por palavra Chave	20
Busca estendida por Palavra Chave a outros Servidores	13
Busca por HASH de arquivo	4

Busca por HASH estendida a Clientes	19
Ambiente de Depuração	22
Bugs	17
Testes	15
Relatórios de Resultados	5

**Total de Horas**            **412**

**Tabela 6 - Horas gastas para produção do protótipo**

#### **4.7.1      AUXÍLIO EXTERNO**

Todos os módulos foram programados integralmente pelo autor deste trabalho com exceção:

- **Pacote Twisted:**

Auxílio indireto sobre referências e demonstrações de uso com código fonte por Rodrigo Senra (Doutorando, UNICAMP)

- **Módulo de Leitura/Escrita:**

Auxílio sobre construção de structs binários por Aloysio Figueredo (Mestre, IMPA)

- **Busca Estendida, Ambiente de Depuração, Decodificador de Pacotes, Bugs:**

Escrita da busca estendida e reescrita do decodificador pacotes, pequenos trechos do ambiente de depuração, correções de bugs em pacotes TCP e UDP, resincronia do pacote Twisted por Cesar Barros (Bacharel, UFRJ)

## 5 ANÁLISE DOS RESULTADOS

### 5.1 PERÍODO EXPERIMENTAL

O período experimental planejado para cada conteúdo foi de vinte e quatro horas, em que cada análise foi composta por 3 arquivos distintos, com formatos de arquivo diferentes, que tivessem no momento da busca (palavra-chave) maior número de fontes. Para cada arquivo escolhido era executada uma estância do protótipo que capturava as informações pertinentes para a análise como é mostrado na figura 27.

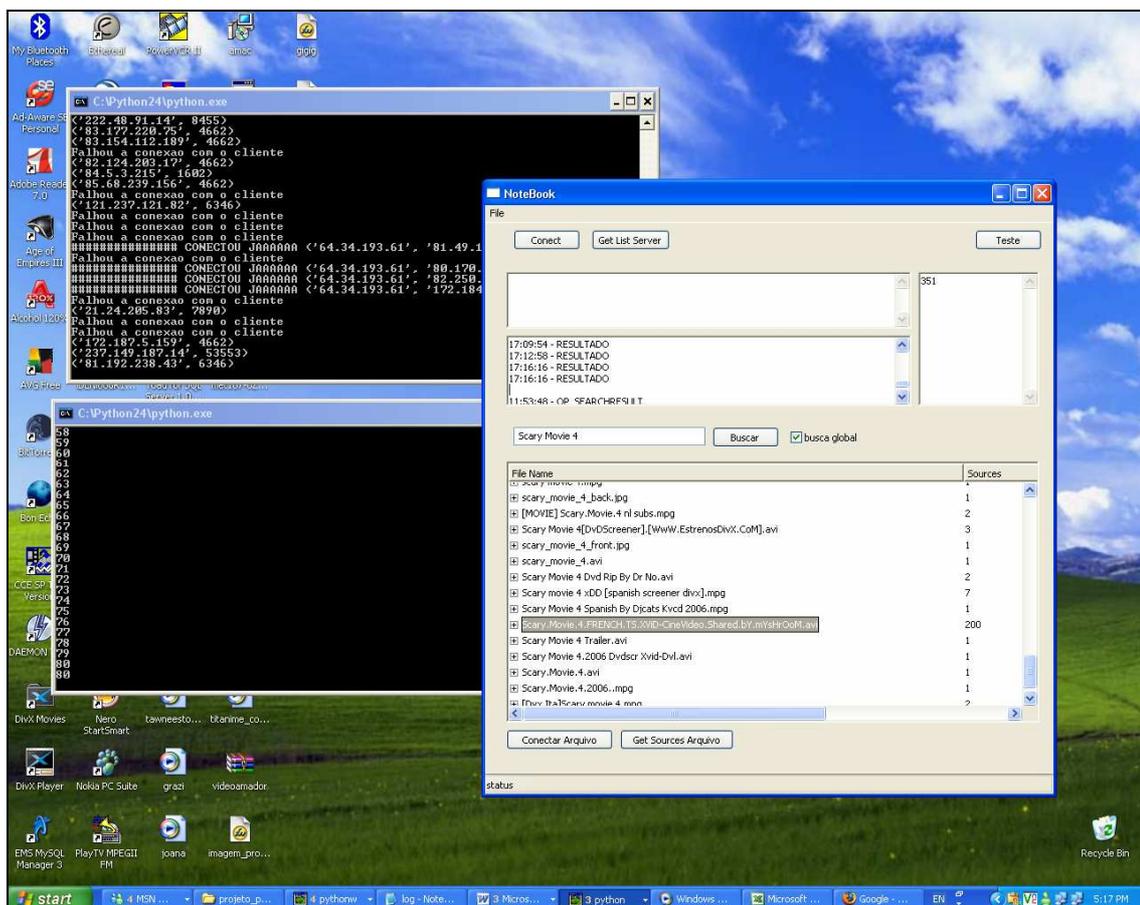


Figura 27 - Tela do protótipo em funcionamento

## 5.2 SEQUÊNCIA DE USO DO PROTÓTIPO:

Para cada conteúdo mensurado utilizou-se a sequência de ações:

- a) Selecionava-se o conteúdo a ser pesquisado por palavras-chave, Ex: “Scary Movie 4”
- b) Capturavam-se os três identificadores únicos (FILE ID) dos arquivos alvos na rede P2P que juntos vão compor o conteúdo P2P metrificado, como mostra a figura 28.
- c) Executava-se uma instância do protótipo para cada File ID
- d) O protótipo gerava dados estatísticos num arquivo tipo XML.

File Name	Size	Availability	Complete Sources	Type	File ID
Scary.Movie.4.FRENCH.TS.XVID-CineVideo.Shared.by.mYshrOoM.avi	693.07 MB	76	78%	Video	5094740822DFECABB79034F9C750E7EC
Scary.Movie.4.TELESYNC.XVID-PUKKA.[oslonet.net].avi	700.69 MB	40	75%	Video	253380D53BF06991ACF530C1059BB00E
Divx Ita Scary Movie 4 2006.avi	704.84 MB	27	70%	Video	423E918140D2EA56AAFE1C177DB5D802
Scary Movie 4 .avi	699.29 MB	25	60%	Video	F62F38028CB5428DAE1C2ABCD0EEBEA
Scary Movie 4.avi	679.71 MB	24	79%	Video	E9AFAD03E608588DD0E34E46FBB61403
Scary Movie 4 Spanish Version Cvcd Divx.mpg	787.58 MB	18	33%	Video	0B9D62E326B018828FF68253A89EDAD7
Scary Movie 4.2006 Italian Md Ts Xvid-Box.avi	699.36 MB	14	71%	Video	486057834689526AA3639EE2D7E63B32
Scary.Movie.4.by.Sbw-Crew.for.www.emuleboard.saarland.6x.to.mpg	794.59 MB	11	81%	Video	A879EEF9455E6AD6384C33ED9B72F1EC
Scary.Movie.4.REP.RSVCD.by.Sbw-Crew.powered.for.emuleboard-sa...	801.76 MB	9	66%	Video	261ED7453FD235D36E25D024C1A9318A

**Figure 28** - Lista de arquivos com seus respectivos FILE ID

## 5.3 ANÁLISES EFETUADAS

O período de testes abaixo foi efetuado do dia 11/06/2006 pelo período de 24 horas com um conteúdo de cada tipo, conforme apresentado na tabela 7.

OBS: A coluna usuários únicos contabilizou apenas os IPs distintos que contivessem um arquivo que remetesse ao conteúdo pesquisado.

Tipo	Nome do Conteúdo	Intérprete / Estúdio / Gravadora	Usuários Únicos
Música Internacional	What's Going On	U2	10028
Música Brasileira	Abalou	Ivete Sangalo	524
Filme Brasileiro	Dois Filhos de Francisco	Globo Filmes e Columbia TriStar	640
Filme Internacional	Scary Movie 4	Miramax	2841
Livro Brasileiro	Onze Minutos (BR)	Paulo Coelho	282
Livro Internacional	Da Vinci Code (EN)	Dan Brown	352

#### **5.4 ANÁLISE DOS RESULTADOS**

Os resultados apresentados na tabela 7 demonstram o quanto em média um determinado conteúdo esteve em trânsito por usuários. Ou seja, quantos usuários únicos trafegaram arquivos com o conteúdo pesquisado durante a varredura do protótipo na rede P2P Edonkey.

Para a interpretação dos dados levaram-se em conta as premissas:

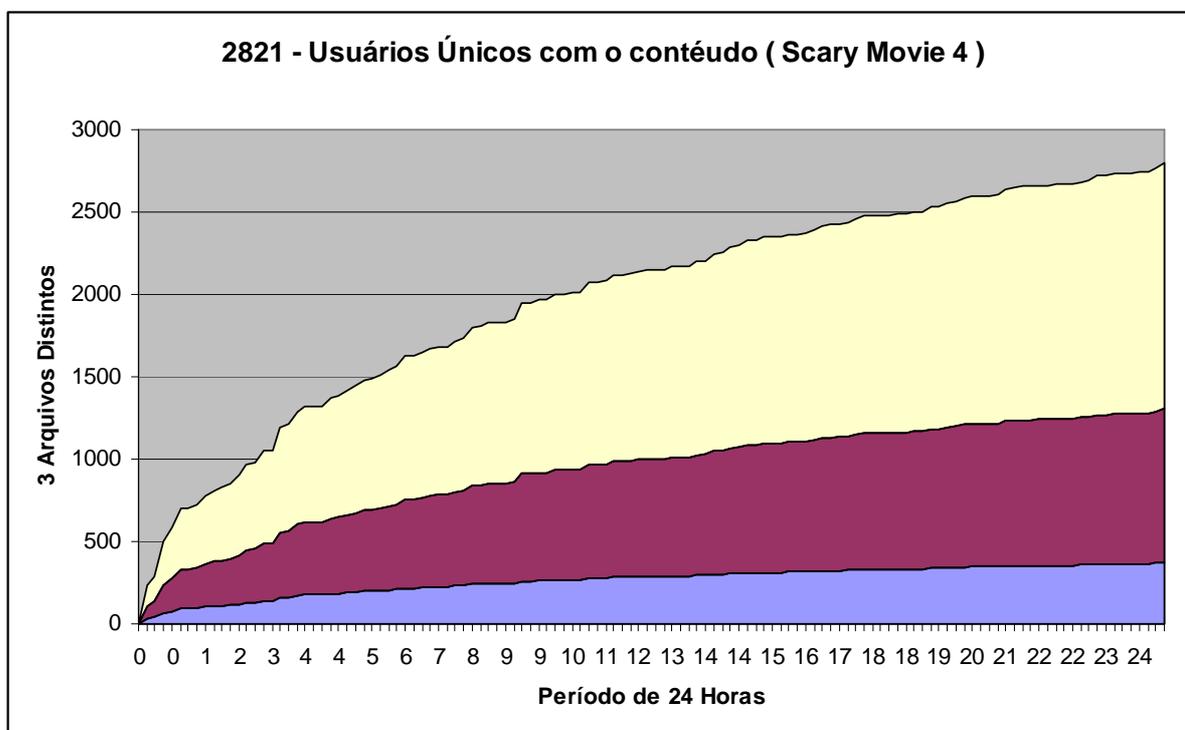
- a) Só foram considerados os usuários que tinham ao menos 20% do conteúdo.
- b) A curva de crescimento do número de usuários dependeu diretamente do interesse pelo conteúdo no dia analisado.
- c) Por existirem muitas variações do conteúdo dentro das redes P2P, foram consideradas apenas as três maiores variações.
- d) A busca inicial se deu por palavras-chave.
- e) O protótipo não garante a leitura de todos os usuários.

#### **5.5 ESTIMATIVA DE INTERESSE DO PÚBLICO**

Para exemplificar o caso, com base nas premissas anteriores foi escolhido o filme internacional “Scary Movie 4”, lembrando que a análise poderia ter sido gerada com qualquer outro conteúdo da tabela 12.

O protótipo conseguiu contabilizar 2821 usuários pelo período de 24 horas, com base em 3 arquivos distintos com o mesmo filme. Sendo assim, é possível chegar a conclusão de que ao menos 2821 usuários tiveram interesse direto neste filme nas vinte e quatro horas analisadas, como pode ser visto na figura 21. As qualidades dos três arquivos eram distintas, sendo a

primeira filmada no cinema, a segunda retirada de rolos de filme das salas de exibição e a terceira com qualidade idêntica a de um DVD com origem desconhecida.



**Figura 29** – Usuários únicos com o filme “Scary Movie 4”

Vale ressaltar que, na data em que foi efetuada esta leitura, o filme ainda estava em apresentação nos cinemas no Brasil e no mundo. Sua versão em DVD só será lançada a partir de 15 de agosto de 2006. Na figura 30 pode ser vista a caixa do DVD que ainda será comercializado.



**Figura 30** – Caixa do Filme “Scary Movie 4”

## **5.6 ESTIMATIVA DE PERDAS FINANCEIRAS:**

Com base no interesse direto dos usuários, uma vez que eles estão baixando o arquivo com o conteúdo do filme “Scary Movie 4”, seria possível simular o valor de uma possível perda financeira direta por parte do estúdio que produziu o longa metragem. Porém para tal previsão eram necessárias que algumas premissas, como:

- a) Conhecer todos os custos envolvidos na produção do lançamento do filme analisado. Como roteiro, filmagem, marketing, etc.
- b) O levantamento de informações de que se todos os usuários que baixassem o conteúdo ilicitamente (ou sem pagar) na rede do Edonkey, teriam interesse de comprá-lo, mesmo que houvesse a disponibilidade para compra física ou virtual.
- c) Se os ganhos financeiros provenientes da produção cinematográfica seria mais alto, caso o filme não fosse trafegado ilicitamente pela rede P2P.
- d) Se de alguma forma pudesse ser comprovada a relações de uma possível queda na bilheteria dos cinemas Brasileiros pelo filme estar disponível com qualidade de DVD na internet antes de ser exibido oficialmente nas salas de exibição.

## 5.7 OUTROS DADOS PERTINENTES LEVANTADOS

Outros dados foram levantados durante o período de testes, como o tamanho médio do arquivo, a qualidade e a quantidade de variações do conteúdo na rede P2P.

<b>Música</b>	
Tamanho	Média de 6 megas, facilitando sua circulação pelas redes P2P
Qualidade	Alta qualidade, em media 128 Kpbs (equivalente ao de um CD)
Quantidade	Cada conteúdo pesquisado tinha em média 16 variações
<b>Filmes</b>	
Tamanho	Média de 700 megas, lento até mesmo para a maioria das conexões de banda larga
Qualidade	Muito variável, de DVD a filmagens da tela do cinema
Quantidade	Cada conteúdo pesquisado tinha em média 5 variações
<b>Livro</b>	
Tamanho	Média de 2,5 megas, muito rápido para circular pelas redes P2P
Qualidade	Em geral muito boa, de arquivos .DOC até PDF permitindo uma impressão perfeita
Quantidade	Em media 10 variações (por conta das traduções para cada língua)

**Tabela 8** – Outros Dados

## **6 CONCLUSÕES**

É urgente uma mudança no modelo de negócio das grandes empresas do segmento do entretenimento. As perdas se acumulam, ano após ano. No mercado fonográfico as lojas de discos já não lotam mais num lançamento. O mercado cinematográfico, esbarrou na facilidade do suporte (DVD) ser clonado e de seus filmes estarem antes na Internet do que nas salas de cinema. Alheia a isso tudo, a geração Internet avança a passos largos ameaçando organizações que por décadas ditaram padrões de consumo e comportamento.

Esta geração nasceu conectada e só veio acelerar a transição do mundo analógico para o mundo digital onde é incentivada por uma cultura livre de barreiras geográficas e ideológica; onde os próprios usuários criam suas próprias leis, alheias à sociedade convencional, criando uma sociedade sem paralelos, sem raça, nação ou credo; onde o conceito de propriedade não existe, e tudo pertence à comunidade através do compartilhamento.

As redes P2P, apesar de todos os esforços judiciais, não apresentam nenhuma queda de crescimento, e a cada dia que passa novas redes surgem e novos usuários conhecem o mundo onde é normal não pagar nada para conseguir músicas, livros ou filmes.

A indústria do entretenimento não está de braços cruzados com a gratuidade de seus conteúdos na grande rede. Ela vem iniciando a modificação de seus modelos de negócio, mesmo que de forma embrionária. Um bom exemplo desta mudança de paradigma, é o sucesso da loja virtual de músicas da Apple (iTunes) que já comercializa faixas de música via Internet, além de comercialização de conteúdos para celulares que já representam até 6% do faturamento das empresas do segmento.

Os números apresentados neste trabalho demonstraram que é possível mensurar separadamente quanto cada conteúdo é distribuído por uma rede P2P. E que profissionais de marketing e gestão de estúdios, gravadoras ou editoras, em posse desses números, podem vir a desenvolver estratégias na comercialização de seus conteúdos online, ou ao menos tentarem reverter em venda parte do que é trafegado ilicitamente. Estes dados, porém, podem ser poderosos instrumentos para conhecer um pouco do perfil deste público, sobre o que o interessa, de que forma gosta de recebê-lo, etc.

A pirataria digital, mesmo que sem comercialização, hoje é apenas um reflexo dos novos tempos e não representa ainda a maior perda financeira desta indústria. De cada três unidades comercializadas (CD, DVD, livro) no mundo, uma é pirata.

Creio que incentivar a criatividade nos canais de distribuição de conteúdo e a disseminação da cultura digital nas empresas do segmento possam fazer com que essas empresas venham a reagir mais rapidamente a mudanças culturais do público consumidor de maneira a alcançar novamente os patamares de faturamento de outrora.

## **6.1 RECOMENDAÇÃO PARA TRABALHOS FUTUROS**

As recomendações serão feitas em duas partes, a primeira com relação ao protótipo e a segunda com pesquisas que venham a tornar mais úteis os números obtidos pelo protótipo para utilizações futuras.

### **6.1.1 PROTÓTIPO:**

Adicionar ao protótipo:

- Cálculo da profundidade de varredura da árvore de usuários
- Identificar a região geográfica de origem dos usuários
- Corrigir anomalias relacionadas à herança da lista de usuários
- Gerar relatórios em tempo real
- Adicionar outras redes P2P, como Kazaa e Bittorrent

### **6.1.2 PESQUISAS**

**Propor as seguintes pesquisas:**

- Qual a porcentagem dos usuários das redes P2P reverteria seu interesse em compra, caso o mesmo conteúdo fosse fornecido de forma legal?
- Qual a faixa etária e a realidade sócio-econômica dos usuários das redes P2P?
- Após disponibilizado, quanto tempo um conteúdo demora a alcançar seu cume de interesse?



## 7 REFERÊNCIAS BIBLIOGRÁFICAS

[http\_1] ABPD - Justiça julga o Kaaza - Setembro 2005

[http://www.abpd.org.br/noticias\\_internas.asp?noticia=114](http://www.abpd.org.br/noticias_internas.asp?noticia=114)

[http\_2] IFPI - The Newsletter of the Internacional Recording Industry - Outubro de 2005

<http://www.ifpi.org/site-content/library/newsletter15.pdf>

[http\_3] IDGNow - Filesharing: pirataria ou novo conceito global? - 27/12/2005

<http://idgnow.uol.com.br/AdPortalv5/DiversaoInterna.aspx?GUID=902A3D60-11C2-4399-8C8F-993D1E1CCCF3&ChannelID=2000010>

[http\_4] RIAA Brings New Round Of Lawsuits Online Music Thieves - 15/12/2005

<http://www.riaa.com/news/newsletter/121505.asp>

[http\_5] Combate a Pirataria na Internet – Outubro de 2005

[http://www.apdif.org.br/pirataria\\_dados\\_internet.asp](http://www.apdif.org.br/pirataria_dados_internet.asp)

[http\_6] RNP - GT Computação Colaborativa (P2P) – Dezembro 2004

<http://www.rnp.br/pd/gts2003-2004/gt-p2p.html>

[http\_7] Trocar música é ilegal no Brasil. Disponível em:

<[http://www.link.estadao.com.br/index.cfm?id\\_conteudo=3361](http://www.link.estadao.com.br/index.cfm?id_conteudo=3361)>. Data de acesso: 11 de Abril de 2005.

[http\_8] Understanding DVD-Audio, A Sonic White Paper

[http://patches.sonic.com/pdf/white-papers/wp\\_dvd\\_audio.pdf](http://patches.sonic.com/pdf/white-papers/wp_dvd_audio.pdf)

[http\_9] DVD

<http://www.dvdforum.org/forum.shtml>

[http\_10] VHS

<http://pt.wikipedia.org/wiki/VHS>

[http\_11] Blu-ray Disc Association

<http://www.blu-raydisc.com/>

[http\_12] MP3: MPEG Audio Layer-3

<http://www.iis.fraunhofer.de/amm/techinf/layer3/>

[http\_13] Associação Brasileira de Empresas de Disco

[http://www.abpd.org.br/pirataria\\_mundo.asp](http://www.abpd.org.br/pirataria_mundo.asp)

[http\_14] Recording Industry Association of America - Trade group that represents the US recording industry. Its mission is to foster a business and legal

<http://www.riaa.com>

[http\_15]APDIF

<http://www.socinpro.org.br/piratinternet.htm>

NELSON, Minar. (2001) Distributed Systems Topologies. **Peer-to-Peer and Web Services Conference**.

KOENEN, R. (1998). “MPEG4 Overview”. **ISO/IEC**, Roma.

KAU W., (1995) “Digital Image Compression: Algorithms and Standards, Kluwer Academic Pub.”

TANEMBAUM, A.S. (2003) “Computer Networks”. **Prentice Hall** . New Jersey – 4a. edição.

RIFKIN, Jeremy. (2000) “A Era do Acesso”. **Makron Books**. São Paulo - 2ª edição.

TAPSCOTT, D. (1999) “Geração Digital”. **Makron Books** – São Paulo.

TAPSCOTT, D. (1998) “Plano de Ação para uma Economia Digital”. **Makron Books**. São Paulo.

CASTELLS, M. (2003) “Galáxia da Internet: reflexões sobre a Internet, os negócios e a sociedade”. **Jorge Zahar**. Rio de Janeiro.

CASTELLS, Manuel. “A sociedade em rede”. **Paz e Terra** . São Paulo.

DIAS, M. (2000) “Os Donos da Voz. Indústria fonográfica brasileira e mundialização da cultura”. **Boitempo Editorial**. São Paulo.

ADORNO, W.(1983) “Idéias para a sociologia da música”. São Paulo - 2ª edição.

BANDEIRA, Messias(1999). “Underground digital no Brasil: a música nas trincheiras do ciberespaço”. **Dissertação de Mestrado**. Facom/UFBA.

LÉVY, Pierre. (1999) “Cibercultura” - **Editora 34** - São Paulo.

SCHOLLMEIER, Rüdiger.(2002) “A Definition of Peer-to-Peer Networking for the Classification of Peer-to-Peer Architectures and Applications”, In: **IEEE Internet Computing**.

ABBATE, Janet.(1999) “Inventing the Internet”. **MIT Press**, Cambridge.

Alfred, D. e Chandler, Jr. (1977) “The Visible Hand: The Managerial Revolution in American Business”. **Harvard University Press**, Cambridge.

ORAM, Andy (2001). “Peer-to-Peer:Harnessing the Power of Disruptive Technologies”. **O'Reilly Media**,

COULORIS, G; T. Kindberg, J. Dollimore, (2005) “Distributed Systems - Concepts and Design”. **Addison Wesley** - 4º Edição.

TANENBAUM, Andrew.(1997) “Redes de computadores”. **Campus** - Rio de Janeiro, 3º Edição.

ROWSTRON, A.; Druschel, P (2001). “Storage management and caching in PAST, a large-scale, persistent peer-to-peer storage utility”, **18-th ACM Symposium on Operating Systems Principles**.

DRUSCHEL, P; Rowstron, A.(2001) “Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems” **IFIP/ACM Internacional Conference Distribuite Systems Platforms**.

STEFAN, Saroiu; Krishna Gummadi e Steve Gribble. (2002) “A measurement study of peer-to-peer le sharing systems”. **In Proceedings of Multimedia Conferencing and Networking**, San Jose.

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)