

# Uma Base de Sinais Pública Para Ensino e Pesquisa em Telecomunicações

Luciana Rêgo, Igor Almeida, Eduardo Medeiros, Lilian Freitas e Aldebaro Klautau

**Resumo**—O presente trabalho descreve o desenvolvimento de uma base pública constituída de arquivos com formas de onda digitalizadas obtidas de sistemas de telecomunicação sem fio e com fio e simulações. São mostradas aplicações dessa base em atividades de pesquisa e ensino, tal como a familiarização com técnicas modernas de modulação: um sinal DMT obtido por um receptor ADSL de um modem comercial pode ser usado para que usuários implementem diversas etapas, desde a demodulação via FFT, até a decodificação de códigos como Reed-Solomon. O trabalho mostra que, além do ensino, a base também é útil em atividades de pesquisa como, por exemplo, classificação de modulação.

**Palavras-Chave**—Ensino de telecomunicações, classificação de modulação, base pública.

**Abstract**—This work presents the development of a public dataset that consists of files storing waveforms representing digitized signals. These signals are digitized from actual wireless and wireline systems or simulated via software. The work describes applications in education and research. For example, the dataset allows the investigation of modern modulation techniques such as DMT: a signal from the ADSL receiver of a commercial modem can be used by users that have the task of implementing several stages, from demodulation using FFT to Reed-Solomon decoding. Besides its use in education, the dataset is also useful for research activities such as modulation classification, for example.

**Keywords**—Education on telecommunications, modulation classification, public dataset.

## I. INTRODUÇÃO

A modernização dos sistemas de telecomunicações causa impacto na metodologia para o ensino das disciplinas da área. Dentre outros desafios, a academia depara-se atualmente com duas demandas conflitantes: o conteúdo dos cursos tende a ser expandido para acomodar as muitas tecnologias que surgiram nos últimos anos e, por outro lado, a duração dos cursos de graduação vem sendo diminuída, tal como no Processo Bolonha da União Européia [1].

Nesse contexto, proporcionar aos discentes a familiarização com aspectos práticos de sistemas de telecomunicações é um aliado importante para o aumento da eficácia do ensino [2]. Universidades com alto poder aquisitivo podem dotar laboratórios de ensino com equipamentos de custo elevado. Contudo, muitas instituições encontram dificuldades para evitar que seus laboratórios fiquem obsoletos. O presente trabalho aborda o desenvolvimento de uma base de dados pública que permite a investigação de diversos aspectos práticos da

transmissão digital sem a exigência de investimentos, tais como a aquisição de equipamentos de laboratório.

A base consiste de arquivos com formas de onda digitalizadas obtidas de sistemas sem fio (*wireless*) e com fio (*wireline*). A mesma abrange não apenas sinais simulados em computadores (aqui chamados “sintetizados”), mas também sinais reais obtidos de *setups* experimentais.

Dentre as motivações para a construção da mesma, encontra-se a popularidade de bases com sinais multimídia disponíveis na Internet. Por exemplo, bases com imagens [3], [4] são usadas intensamente não apenas em ensino como também em pesquisa. Na área de processamento da fala, os corpora distribuídos pelo *Linguistic Data Consortium* [5] também são bastante usados, apesar de não serem gratuitos. Destaca-se também a recente ação do *International Telecommunication Union* (ITU), disponibilizando gratuitamente quinze gigabytes de áudio, imagem e vídeo [6]. Além dos dados chamados multimídia, existem bases de eletrocardiogramas [7] e outros. Especificamente em transmissão digital, alguns sinais de telefonia celular (GSM) foram disponibilizados na Web [8]. Contudo, não se encontra uma base substancial com sinais usados para transmissão digital, nem mesmo dentre os arquivos do ITU.

A base aqui apresentada, chamada *UFPAtelecom* (disponível em [9]), visa preencher essa lacuna. A mesma é utilizada em disciplinas como Transmissão Digital e Processamento Digital de Sinais na UFPA. A base também beneficia a pesquisa em assuntos relacionados, tais como classificação de modulação [10] aplicada a sistemas de rádio cognitivo [11].

Além da contribuição da base em si, o presente trabalho fomenta a troca de experiências acerca do ensino de telecomunicações no âmbito do SBRT, tal como em [12]–[14]. Em [12], um modem acústico em Matlab é apresentado para auxiliar no ensino de telecomunicações. Em [13] e [14], são apresentados dois *softwares* distintos para auxílio no ensino com aplicações práticas.

Este trabalho encontra-se organizado da forma descrita a seguir. As Seções II e III apresentam *setups* experimentais, com apresentação de *hardware* e *software* utilizados, bem como uma descrição dos sinais disponíveis na base de dados considerando os cenários montados para sistemas sem fio e com fio, respectivamente. A Seção IV contém exemplos práticos da aplicação da base de sinais apresentada no ensino de diversas disciplinas relacionadas a telecomunicações, assim como a Seção V, no âmbito da pesquisa. Na Seção VI a organização da base é descrita de forma concisa.

Os autores atuam no Laboratório de Processamento de Sinais (LaPS), CP: 8619, Universidade Federal do Pará (UFPA), Belém - PA, CEP: 66075-110, Brasil. [www.laps.ufpa.br](http://www.laps.ufpa.br). E-mails: {luciana,igor Almeida,eduardom,liliancf,aldebaro}@ufpa.br.

## II. SISTEMAS WIRELESS

Os sinais desta base de dados foram obtidos com o auxílio de vários dispositivos e *softwares*. Para a digitalização de sinais de sistemas wireless utilizou-se a *Universal Software Radio Peripheral* (USRP), um dispositivo de hardware concebido pra ser usado em conjunto com o *software* do Projeto GNURadio [15]. A maioria do projeto da USRP é *open source*, incluindo esquemas e layouts, por exemplo.

A USRP é constituída de uma placa-mãe à qual podem ser conectadas até quatro placas-filhas: duas transmissoras (TX *daughterboard*) e duas receptoras (RX *daughterboard*). Na placa-mãe, há uma interface de comunicação com o computador, a USB 2.0, além de um chip *field programmable gate array* (FPGA) e conversores analógico-digital (ADC) e digital-analógico (DAC). Os slots em que são conectadas as *daughterboards* estão ligados diretamente aos ADCs e DACs, permitindo taxas de 64 MS/s (*samples per second*) e 128 MS/s, respectivamente.

Além dos sinais digitalizados, utilizou-se o *framework* GNURadio isoladamente para obtenção de sinais sintetizados. Isso permite simular canais de comunicação onde os modelos (ruído, distorção, etc.) são fielmente reproduzidos. O GNU-Radio (sem USRP) foi usado com blocos geradores de ruído com distribuição Gaussiana, Laplaciana e ruído impulsivo. Apesar de não refletirem alguns problemas encontrados na prática, os sinais sintetizados via simulação também são úteis. Por exemplo, apenas os sinais wireless sintetizados tiveram controle preciso da razão sinal-ruído (RSR).

Realiza-se a seguir uma breve explanação de como usar o GNURadio, o qual tem uma documentação que pode ser considerada deficiente, como uma boa parte dos projetos de código livre. Ao trabalhar com o GNURadio, deve-se pensar em blocos de processamento de sinais, instanciados e depois conectados em um grafo. Finalmente, o sistema deve ser iniciado para que o processamento seja efetuado. O núcleo do GNURadio é escrito em C++, mas o grafo de processamento é tipicamente descrito em Python, uma linguagem para *scripts*.

```
#fonte , usrp e modulador
from gnuradio import gr, usrp, blks2
class my_message_source (gr.top_block): #grafo
    arquivo->mod->usrp
    def __init__(self):
        gr.top_block.__init__(self)
        self.source = gr.file_source(gr.sizeof_char , #
            tamanho das amostras
            "mensagem.wav", repeat=False)
        mod = blks2.qam8_mod() #modulador QAM
        self.target = u = usrp.sink_c(0)
        #setar freq. intermediaria (IF)
        res = u.tune(tx[0], #canal duc
            flexboard , #daughterboard
            2.5e9) #valor da IF
        self.connect(self.source , mod , self.target) #
            conectar os blocos
    if __name__ == '__main__':
        s = my_message_source()
        s.run()
```

No código em Python acima, ilustrando simplificada a programação usando o *framework*, o conteúdo de um arquivo de som é transmitido via USRP, utilizando modulação QAM-8 e uma portadora de 2.5 GHz. A classe `my_message_source`

TABELA I  
CENÁRIOS USADOS PARA MEDIDAS WIRELESS.

Número de obstáculos	Distância (m)			
	2	6	12	20
0	N0.D2	N0.D6	N0.D12	N0.D20
1	N1.D2	N1.D6	N1.D12	N1.D20
2	N2.D2	N2.D6	N2.D12	N2.D20
3	N3.D2	N3.D6	N3.D12	N3.D20

representa o grafo cujos vértices são o arquivo (origem da informação), o modulador QAM-8 e a USRP (transmissor). No construtor, após instanciar os objetos representativos, a USRP e a *daughterboard* adequada (referenciada pelo objeto `flexboard`, omitida por brevidade) são configuradas para a frequência adequada. O grafo é então executado através da chamada ao método `run()`.

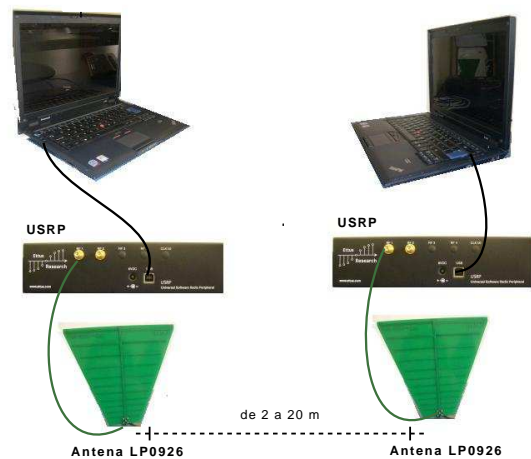


Fig. 1. Ilustração de um setup experimental usando duas USRPs.

A Fig. 1 ilustra um cenário montado para a obtenção de sinais wireless. Como ilustrado, os experimentos práticos são realizados com um conjunto mínimo de equipamentos: notebooks, no caso de mobilidade ser um fator preponderante, e uma USRP por computador, além das antenas. Os cenários construídos tiveram como principais fatores de diferenciação a distância entre as placas de aquisição e o número de obstáculos entre elas. A Tabela I resume as combinações possíveis e mostra os identificadores dos cenários.

Exemplifica-se o processo de obtenção de sinais digitalizados a partir da modulação QAM. O GNURadio não possui atualmente um modem QAM. Assim, foi criado em Python um modulador QAM, para o qual se pode variar diversos parâmetros, tais como o tamanho da constelação e a frequência da portadora, por exemplo. A informação enviada em cada cenário foi a seção de dados de arquivos de áudio do tipo *wav*, de maneira que se pudesse comparar o áudio recebido com o transmitido. No transmissor, um script em Python utilizando o *framework* do GNURadio é responsável pela conexão dos objetos representativos do arquivo de entrada, modulador QAM e USRP. No receptor, um segundo script faz um processo análogo, conectando a USRP a um arquivo de

saída no qual são armazenadas as amostras adquiridas. Porque o GNURadio não fornece um bloco demodulador QAM, pode-se importar os dados para realizar a demodulação no Matlab, por exemplo.

### III. SISTEMAS WIRELINE

Apesar da rápida expansão de sistemas wireless, os sistemas wireline têm importância considerável. Por exemplo, a maioria dos usuários de Internet no Brasil é atendida por linhas digitais de assinante (DSL) [16]. Neste trabalho, os sistemas wireline estão representados por sinais provenientes de modems DSL comerciais não modificados (ADSL [17], ADSL2+ [18]) e também por dados provenientes de um *software* modem chamado JLinmodem, o qual possui implementadas algumas das modulações mais utilizadas por modems *dial-up* (que trabalham na faixa de voz).

O *software* modem *dial-up* foi escrito em Java e é fortemente baseado no Linmodem, escrito em C e desenvolvido no *Linux Soft Modem Project* [19]. O Linmodem implementa padrões como o ITU-V21 [20] e ITU-V23 [21], além de um simulador de linha telefônica e uma interface gráfica para sistemas X11. A transcrição do Linmodem para Java deu-se no intuito de aumentar a legibilidade do código, visto que o original possui poucos comentários, e se beneficiar do paradigma de orientação a objetos. Além disso, os autores estão incorporando ao JLinmodem técnicas de modulação não disponíveis no Linmodem.

Os sinais digitalizados de sistemas wireline (tanto DSL quanto *dial-up*) foram adquiridos utilizando o *setup* experimental representado na Fig. 2. O *setup* constitui-se dos modems DSL do assinante e da central, representados na figura pelo CPEs (*customer premises equipment*) e DSLAM (*digital subscriber line access multiplexer*), respectivamente, cabos telefônicos reais desenrolados (cabos com 30 pares), um gerador de ruído, um injetor de ruído e um gerador de tráfego. Além destes, utilizou-se também o equipamento DSL Xpert da Tracespan, responsável pela aquisição e análise de sinais ADSL e ADSL2+. O Xpert possibilita, além da aquisição do sinal transmitido, a recuperação do *bitstream* completo, incluindo *overhead* de controle da camada física.

Sinais provenientes do *software* modem *dial-up* foram obtidos de maneira semelhante, com a substituição do CPE e DSLAM na Fig. 2 por microcomputadores PC executando os softwares modems JLinmodem. As placas de som dos PCs são usadas para interfaceá-los com o hardware de uma híbrida, desenvolvida pelo grupo, a qual são ligados os pares de fio de cobre.

Os sinais disponíveis na base foram obtidos com as condições de teste listadas na Tabela II, utilizando cabos telefônicos reais, desenrolados, com diâmetro 0.50 mm (24 AWG). Foram variados o tamanho do enlace e também a potência média do ruído AWGN (utilizando o gerador/injetor de ruído).

Para os modems *dial-up* são fornecidos nove arquivos, referentes aos itens presentes na Tabela II. No caso do DSL, através do equipamento Xpert, registrou-se os sinais em diferentes fases da demodulação, sendo elas pré e pós-

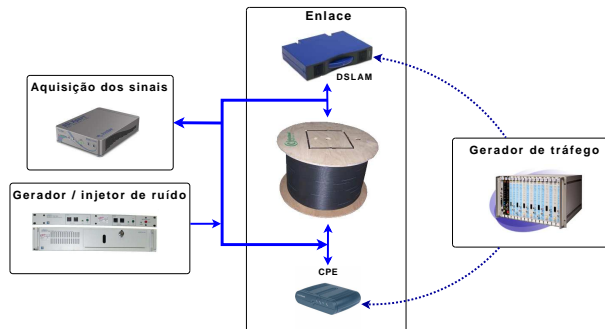


Fig. 2. Setup experimental para aquisição de sinais DSL. Uma configuração semelhante foi adotada para os sinais *dial-up*.

TABELA II  
CENÁRIOS PARA DIGITALIZAÇÃO DE SINAIS DSL E *dial-up*.

Tamanho do enlace (m)	Potência média do ruído AWGN (dBm/Hz)		
	-140	-120	-100
1000	T1_P14	T1_P12	T1_P10
2000	T2_P14	T2_P12	T2_P10
3000	T3_P14	T3_P12	T3_P10

TEQ (*time-domain equalizer*), pós-FFT e pós-FEQ (*frequency-domain equalizer*). O *bitstream*, contendo o *overhead* da camada física, também é fornecido, em três diferentes pontos do receptor: na saída do *de-interleaver*, na saída do *Reed-Solomon encoder* e na saída do *de-scrambler*. Todos os dados são fornecidos para dois serviços distintos, ADSL e ADSL2+. São fornecidos também as alocações de bit utilizadas em cada gravação. Em [22], maiores detalhes acerca desses blocos de processamento encontrados em sistemas DSL são fornecidos.

### IV. APLICAÇÕES NO ENSINO

As subseções seguintes apresentam alguns exemplos de uso dos sinais contidos na base *UFPAtelecom*.

#### A. Análise de Espectrogramas

Espectrogramas (magnitudes de *short-time Fourier transforms*) são apresentados ao aluno para que este identifique modulações, *handshake* e se familiarize com o compromisso entre as resoluções no tempo e frequência. Esta subseção ilustra o estudo do padrão ITU-V21, simulado em *software*.

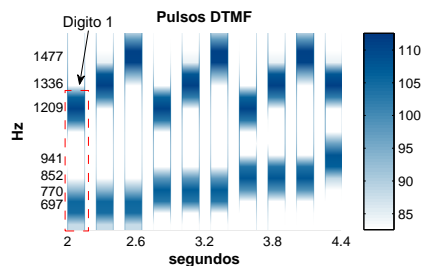


Fig. 3. O modem chamador discar através de pulsos DTMF, onde cada dígito é representado por um par de frequências. Neste caso, o número fictício discado foi 1234567890.

A Fig. 3 mostra o espectrograma de uma parte do sinal transmitido pelo modem chamador, o qual discar através de

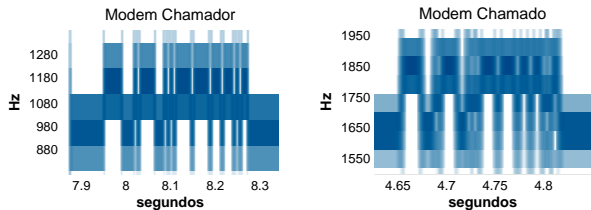


Fig. 4. O modem chamador transmite os bits 1 e 0 em frequências diferentes do modem chamado. No primeiro, as frequências utilizadas são 980 e 1180 Hz, respectivamente. No segundo, as frequências são 1650 e 1850 Hz.

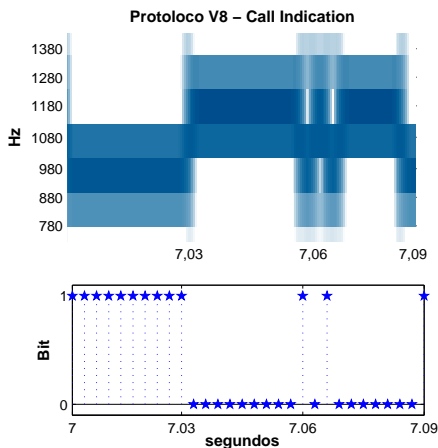


Fig. 5. Os 30 bits do sinal Call Indication do Protocolo V8 enviado pelo modem chamador.

pulsos DTMF (*dual-tone multi-frequency*) e negocia através do protocolo padrão ITU-V8 [23] com o modem chamado. Ao receber a informação de que existe um receptor à espera, este começa a trocar dados através uma simples conexão V21 (modulação FSK). Sabendo quais frequências identificam cada dígito, o discente pode identificar qual número foi discado. Por exemplo, neste caso, os pulsos DTMF correspondem ao número fictício 1234567890, onde se pode observar no início o dígito 1 representado pelo par de frequências 697 e 1209 Hz.

Através da Fig. 4 é possível identificar o padrão ITU-V21 de envio de bits do modem chamador ao modem chamado e vice-versa. O discente pode observar o uso de multiplexação em frequência para que não haja interferência entre os sentidos *downstream* e *upstream*: o modem chamador envia bits identificados pelas frequências 980 Hz para o bit 1 e 1180 Hz para o bit 0, enquanto o chamado representa o bit 0 pela frequência de 1850 Hz e o bit 1 pela frequência de 1650 Hz.

A Fig. 5 ilustra uma parte da negociação V8, utilizada em modems dial-up como o V21 e caracterizada pelo envio de uma cadeia de bits. O *call indication* (CI) é um sinal transmitido a partir do modem chamador para indicar a função geral de comunicação. O CI consiste em uma sequência repetitiva de bits e começa com dez bits 1, os quais precedem cada sequência de informação, seguidos por dez bits de sincronização 000000001, tal como consta no padrão ITU-V8. A próxima informação em CI indica a função de chamada requerida. A sequência 010000001 é interpretada com o primeiro bit 0 revelando o início da informação, os quatro bits seguintes indicam ser uma função de chamada e

os cinco últimos bits apontam que o padrão de modem ITU será determinado posteriormente, além de encerrarem o octeto de informação transmitido com o bit 1.

**B. Ensino de modulações multiportadora**

As técnicas de modulação multiportadora são bastante utilizadas em sistemas de comunicação modernos, sendo os principais exemplos os sistemas DSL, utilizando DMT (*discrete multi-tone*) e vários sistemas wireless baseados em OFDM (*orthogonal frequency division multiplexing*). Daí sua importância em programas educacionais atuais. Os próximos parágrafos apresentam uma breve discussão sobre essas modulações e contextualizam o uso da base.

Os sinais fornecidos pela base, obtidos de sistemas DSL, utilizam DMT como esquema de modulação. Essa técnica baseia-se no uso da transformada de Fourier discreta (DFT) mais prefixo cíclico para particionar o canal, mitigando a interferência intersimbólica, o que permite uma grande melhoria de performance se comparada com sistemas que utilizam uma única portadora.

O grande diferencial da DMT sobre os sistemas OFDM está na capacidade de adaptação às características do canal, através de um processo conhecido como *bitloading*. A RSR é utilizada para definir a quantidade de bits alocados em cada subportadora (subcanal) de acordo com:

$$b_k = \log_2 \left( \frac{RSR_k \cdot \gamma_c}{\Gamma(P_e) \cdot \gamma_m} + 1 \right), \quad (1)$$

onde  $RSR_k$  é a razão sinal-ruído no subcanal  $k$ ,  $\gamma_c$  é o ganho de codificação devido ao uso de códigos corretores de erro,  $\Gamma(P_e)$  é o *gap* para a capacidade de Shannon para uma dada probabilidade de erro de símbolo  $P_e$  e  $\gamma_m$  é uma margem de segurança contra variações do ruído. A Fig. 6 mostra um exemplo de RSR em um sistema ADSL2+ prático e a alocação de bits correspondente.

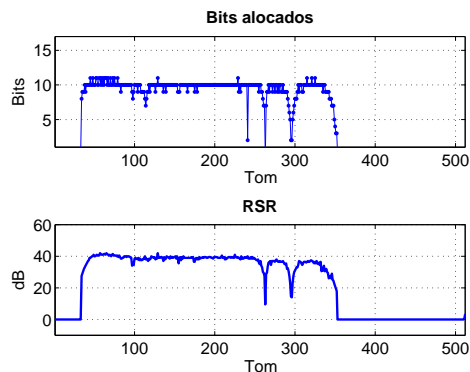


Fig. 6. Alocação de bits e RSR obtidos de um enlace ADSL2+. Mais bits são transmitidos nos subcanais (tons) com melhor RSR.

Os sinais DSL fornecidos na base são acompanhados pelos *bitstreams* transmitidos e pelas alocações de bits utilizadas. Este material pode ser aproveitado para facilitar o aprendizado da modulação em questão e de outros esquemas de modulação multiportadora, enriquecendo as aulas com exemplos práticos. Os sinais gravados podem ainda ser utilizados como auxílio

no ensino e investigação de pontos como a equalização em sistemas multiportadora, ou o uso da DFT para particionamento de canal.

C. Ensino de transmissão wireless

Sinais gerados com o GNURadio e USRP são usados para o estudo de modulação QAM. Algumas características dos sinais obtidos podem ser vistas nos gráficos de magnitude do espectro obtidos via FFT (*Fast Fourier Transform*) da Fig. 7, onde dB FS são os decibéis relativos ao fundo de escala do ADC. Usou-se nesse exemplo uma portadora em 2,5 GHz. Tais gráficos são usados na UFPA para ilustrar a falta de simetria em torno da portadora do sinal QAM (não haver simetria Hermitiana em seu envelope complex). Estuda-se também o aumento na largura de banda utilizada com o aumento da taxa de sinalização. Os aspectos de propagação também são discutidos, com o número de obstáculos no cenário impondo distintas atenuações para diferentes faixas de frequência.

A Fig. 7 compara dois cenários: *visada direta* e *3 obstáculos*, onde é possível observar a atenuação no sinal digitalizado. Discute-se como combater esse problema configurando-se via *software* o ganho dos *programmable gain amplifiers* disponíveis nos conversores em até 20 dB.

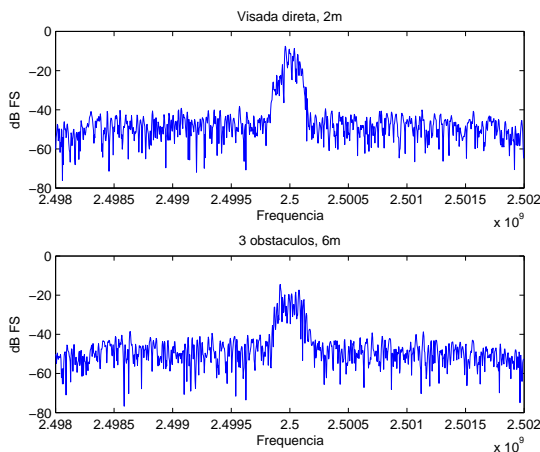


Fig. 7. Exemplos de espectros de modulação QAM com portadora em 2,5 GHz para cenários sem obstáculos e com três obstáculos.

V. APLICAÇÕES NA PESQUISA

A presente seção ilustra como a base UFPAtelecom vem sendo usada em pesquisas na área de classificação de modulação. A classificação automática de modulação (CAM) consiste em identificar o esquema de modulação de um sinal transmitido, com uma alta probabilidade de sucesso e em curto intervalo de observação. Discussões recentes sobre o assunto podem ser encontrados em [10], [11]. A CAM vem sendo utilizada há décadas em sistemas militares e outras aplicações. Recentemente, o interesse por CAM foi renovado pelas pesquisas em rádio cognitivo [24].

A maior parte dos diversos métodos propostos para CAM [25] foram inspirados em técnicas de reconhecimento de padrões [26] e/ou teoria da detecção/estimação [27]. A

grande maioria dos trabalhos na literatura utiliza apenas sinais sintetizados a partir de simulações, ou seja, não usam dados obtidos de sistemas reais. Os modelos utilizados nessas simulações variam em complexidade, incluindo canais AWGN (*additive white gaussian noise*) e canais com desvanecimento. O impacto de erros de temporização, fase da portadora, interferência inter-símbolos e outros também vêm sendo estudados. Contudo, como apontado em [10] (Seção 5), há uma grande dificuldade na comparação dentre os vários métodos propostos. Os motivos são vários e incluem a inexistência de uma base pública de sinais modulados.

O presente trabalho sugere então minimizar este problema com a disponibilização de sinais que usam modulação QAM, FSK, PSK, AM e FM. Sinais com outras modulações devem ser incorporados ao longo do tempo como, por exemplo, OFDM. A base desenvolvida permite então maior facilidade na construção de *benchmarks* para comparação dos métodos, e auxilia na importante tarefa de reprodução de resultados dentre grupos distintos [28]. Como exemplo de aplicação em CAM, apresenta-se alguns resultados preliminares na Tabela III, os quais foram obtidos transmitindo um arquivo de som via USRP, utilizando as modulações: QAM-8, QAM-16, QAM-64, QAM-256, com 500 kbauds e taxa de amostragem 2 MHz. A distância entre as USRPs foi de 2 metros, sem nenhum obstáculos entre elas. Nesse caso, a tarefa é identificar qual a modulação do sinal de teste, para isso são utilizados cinco distintos algoritmos de classificação. Apenas para exemplificar, um conjunto de 350 exemplos de cada modulação foi utilizado tanto para treinar quanto testar os classificadores.

TABELA III  
CLASSIFICAÇÃO DE SINAIS MODULADOS EM QAM-8, QAM-16, QAM-64, QAM-256. PORCENTAGEM DE ACERTO PARA CADA CLASSIFICADOR.

Classificador	QAM-8	QAM-16	QAM-64	QAM-256
NaïveBayes	71,1	68,6	30	30
KNN	100	100	100	100
Árvore J.48	94	92,3	92,3	94,3
SVM	100	100	100	100
Rede Neural	95,4	98	96,3	98

Como de esperar essa situação, o KNN (*k-nearest neighbor*) [29] não teve erros nesse caso. O SVM (*support vector machine*) [30] também apresentou bom desempenho. Detalhes e outros testes, mais realistas, com esses classificadores podem ser encontrados em [31].

Além de CAM, outras aplicações podem se beneficiar do conceito de base de sinais sugerido por este trabalho. Espera-se que pesquisas em outros problemas de rádio cognitivo possam também utilizar a base UFPAtelecom na validação de algoritmos e técnicas.

VI. ORGANIZAÇÃO DA BASE

Para os sinais sintetizados, gerou-se versões de diferentes RSR, entre -10 e 15 dB, além do sinal sem ruído. A Tabela IV descreve o método de digitalização ou síntese dos sinais gerados por cada fonte utilizada na construção desta base.

É importante haver um padrão de formatação dos arquivos e diretórios contidos na base para que os sinais sejam facilmente

TABELA IV  
 FONTE DOS SINAIS DIGITALIZADOS E SINTETIZADOS

Fonte de Sinais	Descrição da Obtenção dos Sinais
GNURadio	Simulação em <i>software</i>
USRP	GNURadio em conjunto com a placa USRP
JLinmodem	Simulação em <i>software</i>
DialupSoftmodem	Jlinmodem utilizado em cenários práticos
DSLXpert	Modems comerciais em cenários práticos

encontrados e manipulados. Com essa finalidade, todos os arquivos são armazenados no formato binário, com o objetivo de reduzir espaço em disco, na forma de variáveis *float* (4 bytes) caso números reais ou inteiros de 16 ou 32 bits. Quando os sinais são complexos (exemplo: envelope complexo de QAM), a parte real vem seguida da parte imaginária. A ordem dos dados é *big-endian* (byte mais significativo em primeiro). Os arquivos não possuem cabeçalho (*headerless*) e são identificados com a extensão *raw*. Como consequência, existe a disponibilidade de *softwares* para lidar com tais arquivos nesta base de sinais.

Cada diretório contém um arquivo texto (*ASCII*) que fornece maiores detalhes dos arquivos internos, tais como frequência de amostragem, número de bits por amostra, condição da digitalização e simulação. O nome do diretório informa o sistema e o cenário de digitalização. O nome do arquivo contém a informação da modulação do sinal e do software / equipamento que o sintetizou ou digitalizou. Os arquivos que contiverem sinais com controle de RSR são indicados com o valor em dB de tal razão. Por exemplo, o diretório nomeado como *less\_N0\_D2* possui arquivos de sistema wireless digitalizados a partir do cenário N0\_D2, descrito na Tabela I. O arquivo nomeado como *gnuradio\_qam16\_15dB* contém um sinal QAM com 16 símbolos na constelação, gerado por blocos do GNURadio com RSR igual a 15 dB.

## VII. CONCLUSÕES

Foi descrita uma base pública, consistindo de arquivos com formas de onda digitalizadas obtidas de sistemas de telecomunicação wireless e wireline. Foram mostradas aplicações da mesma em atividades de ensino e pesquisa.

Vários aspectos de importância prática podem ser explorados com o auxílio dessa base. Por exemplo, o compromisso entre resolução no tempo e em frequência pode ser explorado através da análise de sinais FSK via espectrogramas, com a tarefa de identificar visualmente os bits. A base permite também a investigação de técnicas modernas de modulação, como DMT. Por exemplo, um sinal ADSL pode ser usado para a tarefa de treinar a etapa de demodulação DMT via FFT.

Além do uso em atividades de ensino, planeja-se que a base venha a ser útil também em atividades de pesquisa. Constatou-se recentemente que as pesquisas cujos resultados podem ser reproduzidos causam maior impacto [28]. Contudo, isso exige a disponibilidade de bases públicas e organizadas. O presente trabalho é uma primeira ação nesse sentido. Almeja-se expandir a base e torná-la referência em trabalhos de classificação de modulação, por exemplo.

## REFERÊNCIAS

- [1] "Processo de Bolonha: fazer convergir os sistemas de ensino superior," Disponível em: <http://europa.eu/scadplus/leg/pt/cha/c11088.htm>.
- [2] G. Vejarano and F.G. Guerrero, "A constructivist simulation-based methodology for teaching mobile communications," *IEEE Transactions on Education*, vol. 51, no. 4, pp. 468–475, Nov. 2008.
- [3] F. Andres, K. Ono, A. Makinouchi, and K. Kaneko, "High performance multimedia database system support for image processing," in *Database and Expert Systems Applications, 1997. Proceedings., Eighth International Workshop on*, Sep 1997, pp. 761–766.
- [4] "The USC-SIPI Image Database," Disponível em: <http://sipi.usc.edu/database/>.
- [5] "Linguistic Data Consortium," Disponível em: <http://www ldc.upenn.edu/>.
- [6] "ITU-T Test Signals for Telecommunication Systems Database," Disponível em: <http://www.itu.int/net/itu-t/sigdb/menu.htm>.
- [7] "MIT-BIH Database Distribution," Disponível em: <http://ecg.mit.edu/>.
- [8] "The GSM software project," Disponível em: <http://wiki.thc.org/gsm>.
- [9] "Ufpatelecom," <http://www.laps.ufpa.br/ufpatelecom/>, 2009.
- [10] Dobre O.A., Abdi A., Bar-Ness Y., and Su W., "Survey of automatic modulation classification techniques: Classical approaches and new trends," *IET Communications*, pp. 137–156, 2007.
- [11] Yucek T. and Arslan H., "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys & Tutorials*, pp. 116–113, 2009.
- [12] F. Fernandes, C. Panazio, and D. Zanatta Filho, "Implementação de um software-defined radio com processamento em tempo real usando sinais acústicos e matlab," *Anais do XXVI SBrt*, Set 2007.
- [13] Antonio J.M. Soares et al, "Módulos multimedia para o ensino em sistemas de telecomunicações," *Anais do XVIII SBrt*, Set 2000.
- [14] R. M. Miranda and Luiz R. Mendes, "Ferramenta para ensino das técnicas de compressão de vídeo utilizadas no padrão mpeg-2," *Anais do XXVI SBrt*, Set 2005.
- [15] "GNURadio - the GNU software radio," <http://www.gnu.org/software/gnuradio/>.
- [16] ANATEL, "Relatório anual 2008," .
- [17] ITU-T, "Asymmetric Digital Subscriber Line (ADSL) transceivers," June 1999.
- [18] ITU-T, "Asymmetric Digital Subscriber Line (ADSL) transceivers - Extended bandwidth ADSL2 (ADSL2+)," .
- [19] "A generic Linux Soft Modem," Disponível em: <http://bellard.org/linmodem.html>.
- [20] "V.21: 300 bits per second duplex modem standardized for use in the general switched telephone network," Disponível em: <http://www.gnu.org/software/gnuradio/>.
- [21] "V.23: 600/1200-baud modem standardized for use in the general switched telephone network," Disponível em: <http://www.itu.int/rec/T-REC-V.23-198811-I/en>.
- [22] T. Starr, M. Sorbara, J. M. Cioffi, and P. J. Silverman, *DSL Advances*, Prentice-Hall, 2003.
- [23] "V.8: Procedures for starting sessions of data transmission over the public switched telephone network," Disponível em: <http://www.itu.int/rec/T-REC-V.8-200011-I/en>.
- [24] Simon Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, n. 2, pp. 144–150, Fev. 2005.
- [25] J.H. Reed A. Fehske, J. Gaeddert, "A new approach to signal classification using spectral correlation and neural networks," *DySPAN*, pp. 144–150, Nov 2005.
- [26] R. Duda, P. Hart, and D. Stork, *Pattern classification*, Wiley, 2001.
- [27] H. Van Trees, *Detection, estimation, and modulation theory*, Wiley, 2003.
- [28] P. Vandewalle, J. Kovacevic, and M. Vetterli, "Reproducible research in signal processing," *IEEE Signal Processing Magazine*, vol. 37, pp. 37–47, 2009.
- [29] I. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*, Morgan Kaufmann, 2005.
- [30] S. Haykin and B. V. Veen, *Signals and Systems*, Wiley, 1999.
- [31] Lilian Freitas, "Classificação de modulação para rádio cognitivo," M.S. thesis, UFPA, 2009.