

Andrea Carolina Flores Rodriguez

COMPRESSÃO DE DADOS DE DEMANDA ELÉTRICA EM SMART
METERING

Campinas
2014

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação

Andrea Carolina Flores Rodriguez

COMPRESSÃO DE DADOS DE DEMANDA ELÉTRICA EM SMART METERING

Dissertação de mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos exigidos para a obtenção do título de Mestra em Engenharia Elétrica. Área de concentração: Telecomunicações e Telemática.

Orientador: Gustavo Fraidenraich

Este exemplar corresponde à versão final da dissertação defendida pela aluna, e orientada pelo Prof. Dr. Gustavo Fraidenraich.

Campinas, SP
2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

F663c Flores Rodriguez, Andrea Carolina, 1987-
Compressão de dados de demanda elétrica em *Smart Metering* / Andrea Carolina Flores Rodriguez. – Campinas, SP : [s.n.], 2014.

Orientador: Gustavo Fraidenraich.

Coorientador: Pedro Luis Dias Peres.

Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Medição inteligente. 2. Medidores elétricos. 3. Compressão de dados (Computação). 4. Markov, cadeia de. 5. Entropia (Teoria da informação). I. Fraidenraich, Gustavo, 1975-. II. Peres, Pedro Luis Dias, 1960-. III. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Data compression electricity demand in *Smart Metering*

Palavras-chave em inglês:

Smart metering

Smart meter

Data compression

(Computer) Markov chain

Entropy (Information theory)

Área de concentração: Telecomunicações e

Telemática **Titulação:** Mestra em Engenharia Elétrica

Banca examinadora:

Gustavo Fraidenraich

[Orientador] Moisés Vidal Ribeiro

Dalton Soares Arantes

Data de defesa: 08-08-2014

Programa de Pós-Graduação: Engenharia Elétrica

COMISSÃO JULGADORA - TESE DE MESTRADO

Candidata: Andrea Carolina Flores Rodriguez

Data da Defesa: 8 de agosto de 2014

Título da Tese: " Compressão de Dados de Demanda Elétrica em Smart Metering"

Prof. Dr. Gustavo Fraidenraich (Presidente):

Prof. Dr. Moisés Vidal Ribeiro:

Prof. Dr. Dalton Soares Arantes:

Resumo

A compressão dos dados de consumo residencial de energia elétrica registrados torna-se extremadamente necessária em *Smart Metering*, a fim de resolver o problema de grandes volumes de dados gerados pelos medidores. A principal contribuição desta tese é a proposta de um esquema de representação teórica da informação registrada na forma mais compacta, sugerindo uma forma de atingir o limite fundamental de compressão estabelecido pela entropia da fonte sobre qualquer técnica de compressão disponibilizada no medidor. A proposta consiste na transformação de codificação dos dados, baseado no processamento por segmentação: no tempo em taxas de registros de 1/900 Hz a 1 Hz, e nos valores de consumo residencial de energia elétrica. Este último subdividido em uma compressão por amplitude mudando sua granularidade e compressão dos dados digitais para representar o consumo com o menor número de bits possíveis usando: PCM-Huffman, DPCM-Huffman e codificação de entropia supondo diferentes ordens de distribuição da fonte. O esquema é aplicado sobre dados modelados por cadeias de Markov não homogêneas para as atividades dos membros da casa que influenciam no consumo elétrico e dados reais disponibilizados publicamente. A avaliação do esquema é feita analisando o compromisso da compressão entre as altas taxas de registro, distorção resultante da digitalização dos dados, e exploração da correlação entre amostras consecutivas. Vários exemplos numéricos são apresentados ilustrando a eficiência dos limites de compressão. Os resultados revelam que os melhores esquemas de compressão de dados são encontrados explorando a correlação entre as amostras.

Palavras-chave: *Smart metering*, medidores elétricos, compressão de dados, entropia, cadeias de Markov, teoria de codificação de fonte, PCM, DPCM, Huffman.

Abstract

Data compression of recorded residential electricity consumption becomes extremely necessary on *Smart Metering*, in order to solve the problem of large volumes of data generated by meters. The main contribution of this thesis is to propose a scheme of theoretical representation of recorded information in the most compact form, which suggests a way to reach the fundamental limit of compression set by the entropy of the source, of any compression technique available in the meter. The proposal consists in the transformation of data encoding, based on the processing by segmentation: in time by registration rate from 1/900 Hz to 1 Hz, and in the values of residential electricity consumption. The latter is subdivided into compression: by amplitude changing their regularity, and digital data compression to represent consumption as few bits as possible. It is using PCM-Huffman, DPCM-Huffman and entropy encoding by assuming different orders of the source. The scheme is applied to modeled data by inhomogeneous Markov chains to create the activities of household members that influence electricity consumption, and real data publicly available. The assessment scheme is made by analyzing the trade off of compression between high registration rates, the distortion resulting from the digitization of data, and analyzing the correlation of consecutive samples. Several examples are presented to illustrate the efficiency of the compression limits. The analysis reveals that better data compression schemes can be found by exploring the correlation among the samples.

Smart metering, smart meter, data compression, entropy, Markov chains, coding theory, PCM, DPCM, Huffman.

Sumário

1	Introdução	1
2	Rede Elétrica Inteligente - <i>Smart Grid</i>	4
2.1	Medição de Energia Elétrica	4
2.1.1	AMR à AMI	5
2.2	Smart Grid	6
2.2.1	<i>Smart Grid</i> no Brasil	8
2.3	Resumo do Capítulo	10
3	Caracterização da Demanda Média Residencial e Teoria da Codificação de Fonte	11
3.1	Curva de Carga de um Consumidor Residencial	11
3.2	Modelagem de Fonte de Dados de Consumo de Energia	13
3.3	Modelo de Markov e Membros Ativos de uma Residência	14
3.4	Agregação Temporal	17
3.5	Teoria de Codificação da Fonte de Dados	19
3.5.1	Compressão em Amplitude - Quantização	20
3.6	Resumo do Capítulo	24
4	Limite Máximo de Compressão de Dados do Consumo Elétrico	26
4.1	Modelo de Compressão de Dados de Consumo Elétrico	26
4.2	Modelagem de Demanda de Energia Elétrica Residencial	26
4.2.1	Distribuição Gaussiana na Aleatoriedade do Uso da Carga	29
4.2.2	Modelo Markoviano na Geração de Atividades dos Residentes	30
4.2.3	Validação de Consumo de Energia Elétrica	33
4.3	Avaliação de Compressão	40
4.3.1	Distorção na Compressão da Informação de Consumo	40
4.3.2	Taxa de Registo de Dados de Consumo	40
4.3.3	PCM-Huffman e DPCM-Huffman	44
4.3.4	Ordem de Entropia	46
4.4	Resumo do Capítulo	48

5	Conclusões e Perspectivas	49
5.1	Soluções futuras	49
	Bibliografia	51

DEDICO ESTA TESE AOS MEUS PAIS, MARÍA E VICENTE
A MINHA IRMÃ, CARLA,
E A MEUS IRMÃOS, ANDRES E CARLOS.

Agradecimentos

Agradeço,

A Deus pela força em todos os momentos, especialmente nos mais difíceis, me dar a fé para superar qualquer dificuldade, e me mostrar que a vida é bela.

A toda minha família pelo suporte, especialmente à minha irmã pelo motivação, carinho, e presente companhia, por me fazer sentir que não estou longe dela. Aos meus pais, Maria e Vicente, pelos bons conselhos, confiança, apoio e dedicação constante.

Ao meu orientador, o professor Dr. Gustavo Fraidenraich, infatável disposição, confiança, incentivo em todos os momentos, e da oportunidade de participação neste trabalho.

Às minhas caras amigas: Viviana, Fabíola, Narcleire, por se converter na minha segunda família, por me brindar todo o apoio, carinho e uma agradável convivência. Um especial agradecimento para Alice, pela linda amizade, e os bons momentos que temos compartilhado juntas, vão sempre permanecer na minha memória.

Aos meus grandes amigos: Felipe, Rodolfo, Pâmela, Esdras, pela linda amizade e experiências compartilhadas no diário. Um especial agradecimento para meu caro Juan pelo apreço, companhia, carinho, entusiasmo contagiante, e suas valiosas palavras.

A todos os meus colegas do laboratório que, de alguma forma, contribuíram ao bom andamento deste trabalho.

Aos professores Dr. Moisés Vidal Ribeiro e Dr. Dalton Soares Arantes pela sua colaboração na banca examinadora e ter entregado valiosos comentários e discussões técnicas que aportaram na conclusão deste trabalho.

À CAPES, pelo apoio financeiro concedido durante o programa de mestrado.

À FEEC/UNICAMP pela aceitação e oportunidade de cumprir este sonho.

Lista de Figuras

2.1	Evolução dos tipos de medidores elétricos. (a) Medidor eletromecânico, (b) Medidor eletrônico inteligente ou <i>Smart Meter</i> junto com o dispositivo de visualização de consumo de energia, o qual é instalado dentro da residencia do consumidor. . .	5
2.2	Leitura de medição avançada AMR, com leituras remotas de tipo "drive-by". . .	6
2.3	Infraestrutura de medição avançada AML.	6
2.4	Modelo conceitual do <i>Smart Grid</i> e seus principais domínios. Infraestrutura formada pela integração de um nível de comunicação à rede elétrica atual. . . .	8
3.1	Curva de demanda média representativa do consumidor residencial (<i>Watts</i> $\times$ <i>Tempo</i>), com registros diários em intervalos de $T = 15 \text{ minutos}$, gerando 96 <i>amostras</i> , e consumo de energia diária de 8,84 <i>kWh/dia</i>	12
3.2	Saídas do processo aleatório de consumo elétrico. Cada realização representa um perfil individual de consumo residencial, com registros diários em intervalos de $T = 15 \text{ minutos}$, gerando 96 <i>amostras</i>	13
3.3	Diagrama de estados de uma cadeia de Markov com dois estados ou $m = 1$. Sendo $x_n = i$ e $x_{n+1} = j$, P_{ij} e P_{ji} são as probabilidades de transição de sair de um estado para outro, P_{ii} e P_{jj} são as probabilidades de manter-se no mesmo estado.	15
3.4	Cinquenta exemplos de perfis de ocupação ativos, de uma família de 1 membro em um dia da semana, extraídos a partir dos dados TUS	16
3.5	Perfil de ocupação dos membros ativos em uma residencia de 5 membros. Dados modelados usando matrizes de transição de cadeia de Markov não homogênea, resultantes dos registros TUS.	19
3.6	Curva total de residentes ativos de 123 residencias, e famílias de 1 à 6 membros. Correspondência entre o comportamento das atividades dos membros da casa com a curva de demanda elétrica residencial.	20
3.7	Esquema PCM	21
3.8	Esquema DPCM	22
4.1	Esquema de compressão dos dados de consumo elétrico residencial.	27
4.2	Esquema do processo de modelagem de demanda considerando uma distribuição normal para a aleatoriedade de cada aparelho.	30

4.3	Esquema do processo de modelagem de demanda considerando a participação das atividades dos membros de uma residência.	31
4.4	Dez exemplos de curvas diárias de consumo elétrico residenciais simuladas e a curva correspondente de desvio representando a diversidade de uso dos aparelhos. Resultados usando o primeiro modelo.	34
4.5	Dez exemplos de curvas diárias de consumo elétrico residencial simuladas e a correspondente curva de desvio representando a diversidade de uso dos aparelhos. Resultados usando o segundo modelo.	35
4.6	Dez exemplos de curvas diárias de consumo elétrico residencial reais e a correspondente curva de desvio representando a diversidade de uso dos aparelhos. . . .	36
4.7	Curva de carga representativa do consumidor residencial.	37
4.8	Energia consumida diária em 60 dias para dados simulados e reais.	38
4.9	Distribuição do consumo de potências médias para dados simulados e reais. . . .	39
4.10	Relação sinal-ruído de quantização em PCM e DPCM, para dados de consumo simulados e reais, medidos em intervalos de $T = \{1\text{seg}, 1\text{ min}, 5\text{ min}, 10\text{ min}, 15\text{ min}\}$	41
4.11	Comprimento médio em bits por amostra de medição de consumo diário para os dados simulados e reais, em função dos níveis de quantização e intervalos de registro $T = \{1\text{seg}, 1\text{ min}, 5\text{ min}, 10\text{ min}, 15\text{ min}\}$	43
4.12	Comprimento médio em bits por medição de consumo diário para os dados simulados e reais com 256 níveis de quantização para uma fonte uniformemente distribuída e de primeira ordem, esquemas PCM-Huffman e DPCM-Huffman, em períodos de $T = 1\text{ minuto}$ e 15 minutos	45
4.13	Comprimento médio em bits por amostra para consumo diário em função dos níveis de quantização para diferentes ordens de distribuição da fonte. Dados simulados e reais em $T = 1\text{ minuto}$. Linha contínua representa "Simulado - Modelo 1", linha tracejada "Simulado - Modelo 2" e pontilhada "Real"	47

Lista de Tabelas

2.1	Cidades com projetos de <i>Smart Grid</i>	9
3.1	Distribuição inicial de ocupação ativa para uma família de 5 membros. Distribuição resultante dos registros TUS.	17
3.2	Matriz probabilidade de transição não homogênea de i à j ocupantes ativos $P_{ij}(n)$, para às 03 : 00, 12 : 00 e 19 : 00 horas do dia respectivamente, em uma residência de 5 membros. Matrizes resultantes dos registros TUS.	18
4.1	Especificações de aparelhos domésticos residenciais - Modelo 1	29
4.2	Especificações de aparelhos domésticos residenciais - Modelo 2.	32
4.3	Média e desvio padrão da energia consumida diária dados simulados e reais. . .	36
4.4	Relação sinal-ruído de quantização para PCM e DPCM com 256 níveis de quantização, aplicados aos dados simulados e reais, com medições em períodos de $T = \{1\text{seg}, 1\text{ min}, 5\text{ min}, 10\text{ min}, 15\text{ min}\}$	42
4.5	Comprimento médio em bits por medição de consumo diário para 256 níveis de quantização, aplicados aos dados simulados com medição em períodos $T = \{1\text{seg}, 1\text{ min}, 5\text{ min}, 10\text{ min}, 15\text{ min}\}$	44
4.6	Numero médio de bits totais diários para representar o consumo residencial de uma região de 17 milhões de residentes assumindo 256 níveis de quantização para períodos de $T = \{1\text{ min}, 15\text{ min}\}$	46
4.7	Comprimento médio em bits por medição de consumo diário para 256 níveis de quantização e períodos de 1 minuto para os esquemas PCM-Huffman, DPCM-Huffman e processos de primeira e segunda ordem.	48

Lista de Acrônimos

LCD	<i>Liquid Crystal Display</i> (Tela de Cristal Liquido)
AMR	<i>Automatic Metering Reading</i> (Leitura de Medição Automática)
AMI	<i>Advanced Metering Infrastructure</i> (Infraestrutura de medição avançada)
ICT	<i>Information and Communications Technology</i> (Tecnologias de Informação e Comunicação)
WIFI	<i>Wireless Fidelity</i> (Fidelidade Sem Fio)
WIMAX	<i>Worldwide Interoperability for Microwave Access</i> (Interoperabilidade Mundial para Acesso Micro-Onda)
GPRS	<i>General Packet Radio Service</i>
PLC	Power Line Communication (Comunicação via rede elétrica)
NIST	<i>National Institute of Standards and Technology</i>
CPqD	Centro de Pesquisa de Desenvolvimento em Telecomunicações
ANATEL	Agência Nacional de Telecomunicações
ANEEL	Agência Nacional de Energia Elétrica
SINPHA	Sistema de Informação de Posses e Hábitos de Uso de Aparelhos Elétricos
PROCELINFO	Centro Brasileiro de Informação de Eficiência Energética
DTCV	<i>Discrete-Time/Continuous-Valued</i> (Tempo Discreto/Valores Contínuos)
DTDV	<i>Discrete-Time/Discrete-Valued</i> (Tempo Discreto/Valores Discretos)
HV	<i>High Voltage</i> (Alta Tensão)
UHV	<i>Ultra High Voltage</i> (Ultra Alto Tensão)
MV	<i>Medium Voltage</i> (Média Tensão)
LV	<i>Low Voltage</i> (Baixa Tensão)
TUS	<i>Time-Use Survey</i> (Pesquisa de Uso do Tempo)
A/D	Análogo/Digital
PCM	<i>Pulse Code Modulation</i> (Modulação por Codificação de Pulso)
DPCM	<i>Differential Pulse Code Modulation</i> (Modulação por Codificação de Pulso Diferencial)
SQNR	<i>Signal-to-Quantization-Noise Ratio</i> (Relação Sinal Ruído de Quantização)
REED	<i>Reference Energy Disaggregation Data Set</i> (Conjunto de Dados de Referência para Desagregação de Energia)
UTC	<i>Universal Time Coordinated</i> (Tempo Universal Coordenado)

Lista de Símbolos

T	período de registro da potência média.
E	energia diária consumida.
n	indicador de amostras no tempo.
D_{R_r}	demanda no tempo de cada residência.
r	número do residente.
D_{A_a}	demanda no tempo de cada aparelho.
a	número do aparelho.
D_{m_n}	demanda média residencial no tempo.
D_{max}	maior valor de demanda média no dia.
P_{min}, P_{max}	potências entre os diferentes tipos de lâmpadas comerciais.
L_{lim}	nível de referência de iluminação.
$D_{stand-by}$	demanda dos aparelhos com consumo de potência <i>stand-by</i> .
$D_{continuo}$	demanda dos aparelhos com consumo de potência contínua.
D_{ativa}	demanda dos aparelhos com consumo de potência ativa.
D_{frio}	demanda dos aparelhos com consumo de potência por ciclos de trabalho.
t_n	instante de tempo no dia.
$\dot{t}_n$	instante de tempo no dia resultado da agregação temporal.
D_n	série de dados no tempo.
$\dot{D}_n$	série de dados no tempo resultado da agregação temporal.
$\mathcal{W}$	Alfabeto dos valores de demanda.
$\dot{\mathcal{W}}$	Alfabeto finito dos valores de demanda, resultado da quantização.
w_n	valor de demanda em um instante de tempo.
$\dot{w}_{in}$	valor de demanda quantizada em um instante de tempo.
W_n	variável aleatória de demandas. Função que pode assumir qualquer valor $w_n \in \mathcal{W}$.
$\dot{W}_{in}$	variável aleatória de demandas. Função que pode assumir qualquer valor $\dot{w}_{in} \in \dot{\mathcal{W}}$.
$\mathbf{P}$	matriz de probabilidade de transição de estados.
$P_{x_n x_{n+1}}$	elementos da matriz $\mathbf{P}$.
$Pr(\cdot)$	função de probabilidade.
$num_{ij}(n)$	número de ocorrências favoráveis para a transição entre os estados i e j em n .
$num_i(n)$	número de ocorrências favoráveis para o estado i em n .
X_{n+1}	variável aleatória de estados em $n + 1$.
X_n	variável aleatória de estados em n .
x_n	estado em n .
x_{n+1}	estado em $n + 1$.

$\mathcal{X}$	espaço de estados.
$\mathbb{N}$	conjunto de números naturais.
m	ordem da cadeia ou número de memória da cadeia de Markov.
f	número de membros ou tamanho de uma família.
N	número de regiões de quantização.
$\mathcal{R}_i$	i -ésima região de quantização.
$Q(\cdot)$	quantização uniforme PCM.
e_n	diferença entre a amostra prévia predita e o valor atual.
W'_n	amostra prévia predita.
$\hat{e}_{i,n}$	quantização PCM de e_n .
$\widehat{W}_n$	valor reconstruído.
$\mathcal{D}$	valor de distorção.
$SQNR _{dB}$	relação sinal ruído de quantização, medido em dB.
$E[\cdot]$	média estatística.
b	número de bits.
L	nível de luz.
$\mathcal{L}$	comprimento médio da palavra código.
l_i	comprimento do i -ésimo nível de quantização.
$\mathcal{H}$	taxa de entropia ou limite de compressão.
$\mathcal{H}_1(\cdot)$	entropia de primeira ordem.
$\mathcal{H}_2(\cdot)$	entropia de segunda ordem.
$\mathcal{H}_3(\cdot)$	entropia de terceira ordem.

Introdução

O sistema elétrico de potência foi projetada para operar como uma estrutura vertical consistindo da geração, transmissão, e distribuição, onde a energia elétrica é gerada por um pequeno número de grandes centrais elétricas, transportada em altas voltagens e fornecida em baixas voltagens aos usuários finais usando o sistema de distribuição. A rede elétrica é, possivelmente, o sistema interligado mais extenso e complexo do planeta (Momoh 2012). Durante mais de um século, tem estado intimamente associado ao desenvolvimento tecnológico e tem cumprido com seu propósito perfeitamente. Permitiu entregar a energia produzida, de forma imediata e permanente, com custos reduzidos e razoável confiabilidade. Infelizmente, isso fez com que o sistema elétrico seja o mais esquecido do século XX, deixando a rede está no limite da sua capacidade. Até agora, a única preocupação era a de "manter as luzes acesas", no entanto, nos últimos anos as restrições foram exacerbadas pelo aumento da demanda de energia elétrica, devido ao crescimento populacional e à aquisição massiva de produtos de consumo elétricos, como televisores, computadores e aparelhos de ar condicionado (Gellings 2012, Momoh 2012).

Isso representa um problema que se tornou um foco de investigação fundamental, estimulando a buscar por mecanismos de uso consciente de energia. A partir daí é que surge a atualização da rede elétrica com inteligência, autonomia, eficiência melhorada, controle fácil e alta confiabilidade. O conceito de *Smart Grid* apresenta-se como uma das fortes tendências de modernização do sistema elétrico de potência, que foi criado e introduzido para superar esses desafios, uma vez que as tecnologias de informação e comunicação ou ICTs (do inglês "*Information and Communications Technology*") são agregadas na rede e suportadas por uma infraestrutura de medição avançada ou AMI (do inglês "*Advanced Metering Infrastructure*") (IEEE SmartGrid 2013, Momoh 2012).

O primeiro passo na modernização do sistema elétrico está se apresentando com o desenvolvimento de testes em projetos pilotos para a avaliação do *Smart Grid*. Particularmente, a maior parte do foco da rede inteligente está na distribuição, que será constituída por uma grande variedade de sensores, monitores e medidores inteligentes (ou *smart meters*) espalhados, para permitir uma observabilidade em tempo real e oferecer uma gama variada de benefícios e serviços, tais como faturamento em tempo real, melhor eficiência energética, reconfiguração,

registro instantâneo de falhas, e informação do consumo aos usuários, permitindo a compreensão e conscientização sobre seu consumo.

Embora *Smart Metering* prometa um grande número de vantagens, a substituição dos medidores eletromecânicos pelos eletrônicos, contribuem à geração do chamado "*Big Data*" e a participação na "*Internet of Things*". A quantidade de dados que terão que ser armazenados, processados e transmitidos todos os dias torna-se uma preocupação. Portanto, sistemas de compressão de dados serão muito importantes, a fim de reduzir a utilização do canal de comunicação e requisitos de armazenamento que são limitados nos equipamentos de medição.

Vários algoritmos modernos de compressão de dados têm sido estudados, a fim de se identificar qual deles fornece o melhor resultado para armazenar os dados usando a menor quantidade de recursos. Em vista disso, têm sido exploradas em profundidade várias técnicas de compressão da forma de onda dos sinais elétricos. Assim, esquemas de compressão sem perdas baseados na utilização da transformada rápida de Fourier e codificação preditiva ADPCM com baixa distorção (Mehta & Russell 1989), codificação de Huffman em sinais com modulação delta de alta ordem (Zhang, Bi & Zhao 2009), técnicas de codificação aritmética, Lempel-Ziv, Zip (Kraus, Robiska & Bubla 2009, Lorio & Magnago 2004), e outra variedade de técnicas de compressão sem perdas desenvolvidas em áudio e imagens (Gerek & Ece 2008) foram aplicadas.

No entanto, poucas são as investigações focadas em analisar o desempenho das técnicas de compressão nos dados de consumo residencial de energia. Em (Wijaya, Eberle & Aberer 2013), descreve-se uma técnica de representação simbólica sobre o sinal de dados de consumo no tempo conservando sua forma analítica. Em (Rigwelki, Renner, Reinhardt, Weigel & Turau 2012, Li, Mao, Lai & Qiu 2010), são apresentados vários algoritmos e seus respectivos desempenhos usando dados do medidor inteligente.

Dada a importância do estudo da compressão utilizando técnicas disponíveis para representar a informação registrada pelos medidores residenciais de consumo de energia em uma forma compacta, a presente dissertação foca no limite de compressão de dados usando teoria da informação, que serve de referência para o projeto de compressão sem perda. Este limite teórico refere-se à entropia, que descreve o menor e suficiente comprimento médio em bits por símbolo que garante uma representação sem perdas.

O esquema de compressão proposto consiste na amostragem e na quantização dos valores de consumo residencial de energia elétrica com taxas que variam entre 1/900 Hz e 1 Hz. Em seguida, utiliza-se PCM-Huffman, DPCM-Huffman e codificação de entropia supondo diferentes ordens de distribuição da fonte. O esquema é aplicado para dados de consumo médio residencial reais obtidos em medições feitas em campo (Kolter & Johnson 2011), e simulados usando um modelo por cadeias de Markov não homogêneas. Nas simulações, informações de consumo dos aparelhos, porcentagem de posse e hábitos de uso dos aparelhos elétricos do Brasil (PROCEL & ELETROBRÁS 2005, EFLUL 2013, PROCELINFO 2013, Nunes, Da Costa,

De Oliveira, Goretti, Papa & Sikiti 2010) foram utilizadas para a geração das atividades aleatórias dos residentes que obviamente têm influência na utilização de energia elétrica (Richardson, Thomson & Infield 2008, Richardson & Thomson 2011, Ardakanian, Keshav & Rosenberg 2011).

O texto está organizado em cinco capítulos. No Capítulo 1 é descrito a motivação, os objetivos e a relevância deste trabalho. No Capítulo 2 é abordado uma visualização geral sobre a transição da atual rede elétrica à *Smart Grid*, definição do esquema de rede, principais componentes, vantagens e situação atual no Brasil. No Capítulo 3 são tratadas as diversas formas de modelagem da fonte de dados, caracterização da curva média de consumo residencial, os principais conceitos da teoria de codificação de fonte e as técnicas de compressão de dados sem perda. No Capítulo 4, são apresentados o modelo proposto e os resultados da modelagem da curva de carga simulada e o desempenho da compressão sob diversos cenários. No Capítulo 5, são apresentadas as considerações finais e as sugestões para trabalhos futuros.

Rede Elétrica Inteligente - *Smart Grid*

Neste capítulo são apresentadas uma breve descrição da transição dos sistemas de medição de energia elétrica e as diferenças entre seus principais componentes. Na seguinte seção é exposto a definição e a estrutura geral do *Smart Grid*, suas características, elementos, serviços e diferenças com a rede elétrica atual. Na última seção é descrita a situação do *Smart Grid* no Brasil, que tipo de projetos estão sendo desenvolvidos e quem são seus principais participantes.

2.1 Medição de Energia Elétrica

O medidor de consumo de energia elétrica que os consumidores residenciais dispõem em sua entrada de energia, é o medidor de tipo indução ou eletromecânico como é apresentado na Fig.2.1a, que foi desenvolvido a finais do século XIX. Esse tipo de medidor contém um disco giratório e um indicador do contador mecânico, funcionando através da interação eletromagnética dos fluxos alternados produzidos nas bobinas, criando um movimento de rotação no disco. A velocidade de rotação do disco é função da potência solicitada pela carga, sendo a energia ativa consumida proporcional ao número de voltas do disco. Um ímã permanente exerce uma força de amortecimento no disco, parando seu giro após a energia ter sido removida (Harney 2009, Jardini 1999).

A substituição por um medidor eletrônico de estado sólido foi o primeiro passo na evolução. Este mede o consumo de energia utilizando equipamentos baseados em processadores, os quais permitem digitalizar a tensão e a corrente instantânea com elevada resolução temporal (cujo produto é a potência instantânea), assim como o fator de potência, e potências ativa e reativa. O dado medido de consumo de energia é apresentado na tela de cristal líquido ou LCD (do inglês "*Liquid-Crystal Display*"), e armazenados em intervalos específicos. Pelo fato de não conter um ímã na sua estrutura, as adulterações deixa de ser frequentes e ganhando maior confiabilidade (Harney 2009). Um exemplo deste tipo de medidor é apresentado na Fig. 2.1b.



Figura 2.1: Evolução dos tipos de medidores elétricos. (a) Medidor eletromecânico, (b) Medidor eletrônico inteligente ou *Smart Meter* junto com o dispositivo de visualização de consumo de energia, o qual é instalado dentro da residência do consumidor.

Uma vez que os dados do medidor estão disponibilizados em forma digital, torna-se viável adicionar comunicação remota ao medidor, permitindo as leituras de uma forma automática ou AMR (do inglês "*Automatic Metering Reading*") através de uma via de comunicação, seja de forma manual, a pé ("*walk-by*"), mediante condução ("*drive-by*") ou utilizando algum sistema de rede ("*networked*").

Especificamente, no sistema "*drive-by*" como é apresentada na Fig. 2.2, a empresa distribuidora de energia elétrica envia um veículo com um coletor de dados sem fio que recolhe os dados dos medidores do setor percorrido de uma maneira muito eficiente se comparado ao sistema "*walk-by*" onde as leituras manuais. No sistema "*networked*", os dados do medidor são fornecidos a um coletor fixo, tipicamente localizado em um poste no fim da rua ou vizinhança para posterior envio dos dados à companhia através de um *backbone* de banda larga ou de celular (Harney 2009).

2.1.1 AMR à AMI

A substituição de leituras manuais pelo sistema AMR foi visto simplesmente como uma maneira de reduzir o custo laboral, mas isso foi mudando à medida que a indústria reconheceu que a AMR permitiria às empresas distribuidoras de energia elétrica fornecer mais serviços e benefícios, tais como: tarifação em tempo real para promover uma melhor eficiência energética, reporte instantâneo de falhas, corte e religação, e dados mais precisos para criar um perfil de uso dentro da rede. A AMR passa a ser uma infraestrutura de medição avançada ou AMI (do inglês "*Advanced Metering Infrastructure*") para destacar a evolução de uma simples leitura remota.

O sistema de medição "*networked*" da AMI pode integrar uma variedade de tecnologias de

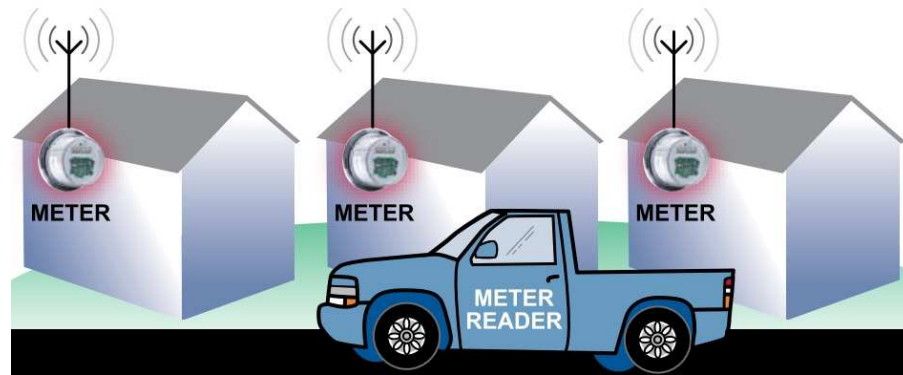


Figura 2.2: Leitura de medição avançada AMR, com leituras remotas de tipo "drive-by".

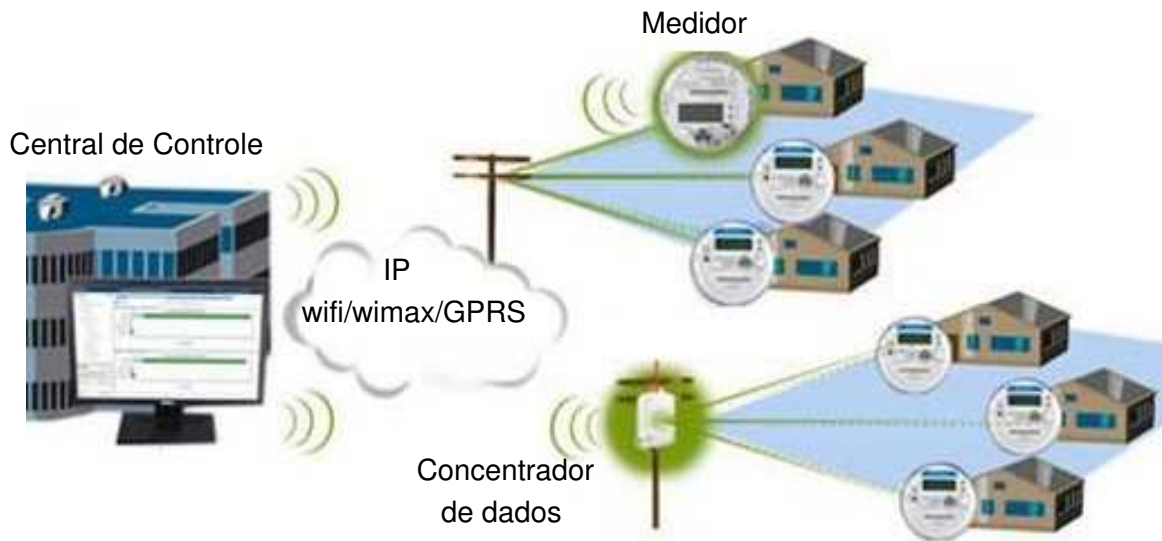


Figura 2.3: Infraestrutura de medição avançada AMI.

informação e comunicação ou ICTs (do inglês "*Information and Communications Technology*"), como é apresentado na Fig.2.3, que vão desde satélites a rádios de baixo custo utilizando bandas não licenciadas e licenciadas (p. ex., Wifi, Wimax e GPRS) e comunicação por fio (p. ex. *power line communication* - PLC), a fim de garantir a comunicação bidirecional entre o usuário e a central de controle. Para isso, o sistema deve usar sensores providos de portas de comunicação de dados, periféricos e inteligência computacional embarcada (Harney 2009, Romano & Sípoli 2013).

2.2 Smart Grid

A atual rede elétrica foi projetada para operar como uma estrutura vertical consistindo da geração, transmissão, e distribuição. Mais especificamente, o gerador fornece energia com uma tensão típica de 13,8 kV e esta tensão passa por uma subestação de geração de modo que a energia é transmitida através de linhas de alta tensão, que vão desde 138 até 1000 kV, referida

como nível de alta tensão ou HV (do inglês "*High Voltage*") e ultra-alta tensão ou UHV (do inglês "*Ultra High Voltage*"), respectivamente. Quando a energia alcança a distribuição, a tensão é diminuída novamente para o nível de média tensão ou MV (do inglês "*Medium Voltage*"), o que caracteriza a rede de distribuição. Normalmente, no Brasil os alimentadores de distribuição primária saem dessa subestação de nível MV em 13,8 kV para diferentes zonas. Transformadores de distribuição estão ligados aos alimentadores primários em muitos pontos, para reduzir o nível de tensão de 13,8 kV a 127 V, 220 V ou 380 V, e assim, energizar os alimentadores de distribuição secundários, que atingem os consumidores finais. Em seguida, o sistema de distribuição elétrica secundária corresponde à alimentadores de baixa tensão ou LV (do inglês "*Low Voltage*") (Momoh 2012). O esquema geral da rede atual é retratado no bloco inferior da Fig. 2.4.

Embora o sistema de fornecimento de energia elétrica tem um uso modesto de sensores, pouca comunicação e controle eletrônico, este pode manter certa confiabilidade e estabilidade. No entanto, as operadoras do sistema estão enfrentando novos desafios com a penetração de fontes de energia renovável, envelhecimento da infraestrutura, rápida mudança de tecnologia, e a variabilidade de participantes no mercado e usuários finais no sistema atual, tornando-se uma preocupação global que requer que a rede elétrica evolua em direção a inteligência, autonomia, eficiência, fácil controle e alta confiabilidade.

Smart Grid se apresenta como uma das fortes tendências de modernização do sistema elétrico, que foi criada e introduzida para superar esses desafios, uma vez que as ICTs são agregadas na rede e suportadas pela AMI (IEEE SmartGrid 2013, Momoh 2012). O modelo conceitual do *Smart Grid* está definido pelo Instituto Nacional de Padrões e Tecnologia ou NIST (do inglês "*National Institute of Standards and Technology*") em diferentes domínios com cada um formado por importantes elementos conectados através de enlaces de comunicações bidirecionais, como é apresentado na Fig. 2.4 (Gellings 2012, IEEE SmartGrid 2013, Momoh 2012).

As principais evoluções previstas com a instalação de AMI na rede elétrica especificamente serão: a substituição dos medidores eletromecânicos pelos digitais inteligentes ou *smart meters*, e equipar os clientes de um *display* que permita ao consumidor controlar melhor seu consumo elétrico, além de ser desnecessário o trabalho dos leituristas da companhia elétrica; a implementação de micro-geração distribuída; e a introdução de sistemas de comunicação que cubra todas as áreas. Isso leva a necessidade da incorporação de novos elementos funcionais como *gateways*, sensores, equipamentos de múltiplas interfaces e padrões que permitam a interoperabilidade entre eles.

Em consequência, permite às companhias monitorar os picos de carga através da participação do consumidor, controlar o consumo de eletricidade para uma geração ótima de potência e balanço energético com a finalidade de combater às perdas e cortes. Desta forma, pode-se aproveitar melhor os processos de distribuição no faturamento, por exemplo variar os preços de consumo (tarifação diferenciada ou tarifa branca) e evitar os picos de consumo, bem como reduzir o custo (Romano & Sípoli 2013, EXAME 2013).

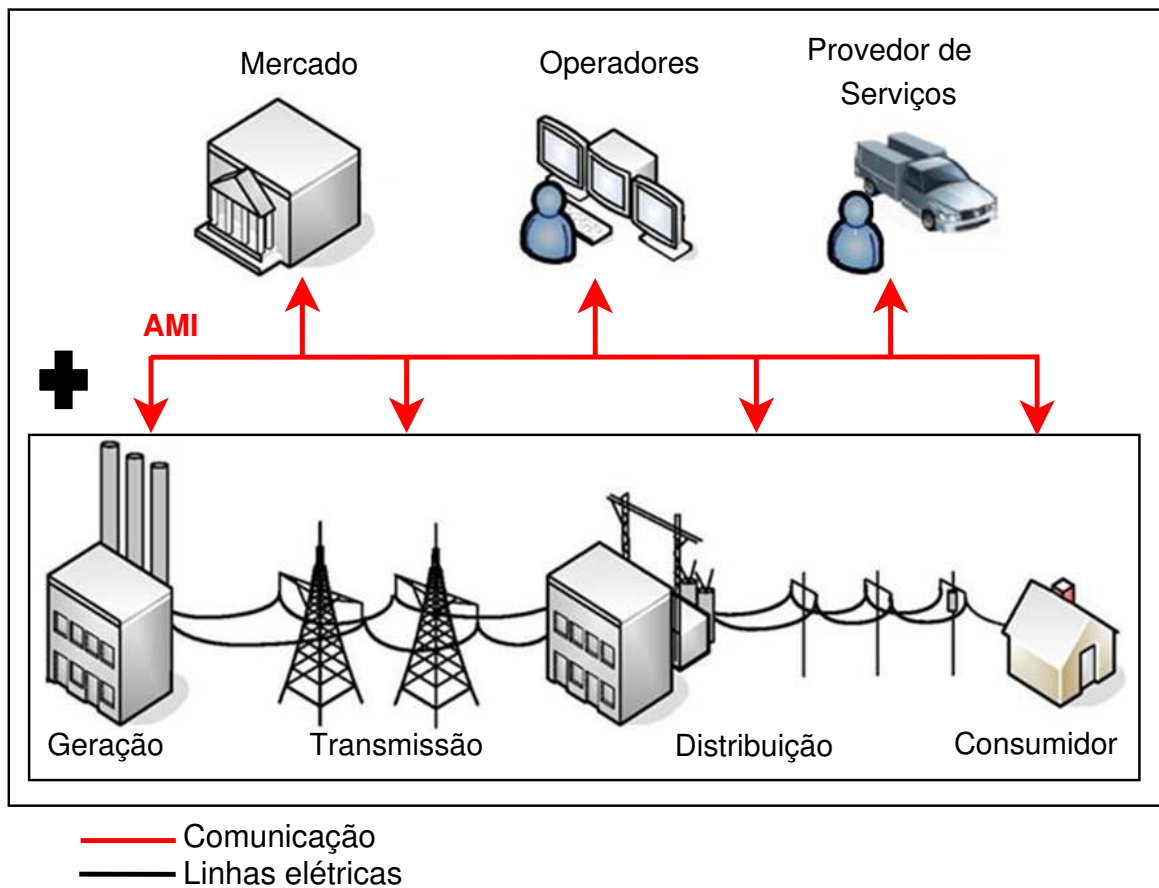


Figura 2.4: Modelo conceitual do *Smart Grid* e seus principais domínios. Infraestrutura formada pela integração de um nível de comunicação à rede elétrica atual.

2.2.1 *Smart Grid* no Brasil

Atualmente a matriz energética do Brasil é baseada principalmente em energias renováveis (hidroelétricas). As distribuidoras de energia elétrica recebem as informações sobre o consumo de energia elétrica dos clientes pelos leituristas apenas uma vez por mês, e também só sabem de um problema no fornecimento se alguém ligar para reclamar.

Nos últimos 25 anos, quase todas as indústrias do mundo ocidental têm-se modernizado com o uso de sensores, sistemas de comunicações e desenvolvimento computacional dentro da rede elétrica, conseguindo enormes melhorias em: produtividade, eficiência, qualidade de produtos e serviços, e meio ambiente. No entanto, ao longo dos últimos dois anos, as principais concessionárias de energia elétrica do Brasil como Cemig, Celpe, Eletrobras, Eletropaulo, Light entre outras, começaram a participar em projetos pilotos sobre a AMI junto com o centro de pesquisa de desenvolvimento em telecomunicações ou CPqD, para a avaliação de soluções de medição avançada e de infraestrutura de telecomunicações bidirecional necessária para suportar essas soluções, e em projetos desenvolvendo a substituição de medidores eletromecânicos por eletrô-

nicos inteligentes, a fim de avaliar a viabilidade do *Smart Grid* em vários municípios do Brasil (EXAME 2013, Romano & Sípoli 2013). A tabela 2.1 lista as cidades onde projetos *Smart Grid* estão sendo desenvolvidos.

Tabela 2.1: Cidades com projetos de *Smart Grid*

Paritins/AM	Eletrobras Amazonas
Fernando de Noronha/PE	Celpe/Neoenergia
Sete Lagoas/ MG	Cemig
Búzios /RJ	Ampla/Endesa
Barueri/SP	AES Eletropaulo
São Paulo / SP	AES Eletropaulo
Rio de Janeiro/RJ	Light
São Luiz do Paraitinga / SP	Elektro
Aparecida/SP	EDP Bandeirantes

O teste pioneiro foi feito pela concessionária Portuguesa EDP com a troca de seus 13000 medidores em Aparecida desde 2010. Em Minas Gerais, a concessionária Cemig finalizou em abril do presente ano a troca de equipamentos para 8000 clientes na região de Sete Lagoas. A AES Eletropaulo, a maior concessionária do país em faturamento e consumo de energia, com 6,5 milhões de consumidores deverá equipar todos os 72000 clientes de Barueri com medidores inteligentes até 2015, para depois expandi-lo para toda a sua área de concessão (EXAME 2013).

Isso quer dizer, que com a adoção do *Smart Grid*, todos os usuários que ganharam o medidor digital inteligente, cujo dispositivo é apresentado na Fig.2.1b, também passarão a ter em mãos uma ferramenta para gerenciar o seu consumo de energia. Além disso, a forte tendência de crescimento deste tipo de usuários, significa uma grande quantidade de dados fluindo pela rede, que terão que ser coletados, transmitidos e processados todos os dias. Assim é como nasce o enfoque sobre o estudo da compressão de dados das leituras de consumo de energia elétrica detalhado neste trabalho.

A modernização da rede elétrica do Brasil, iniciando com a substituição dos medidores, torna-se um assunto importante. Em 2012, cerca de 4% da energia distribuída foi perdida devido perdas nas técnicas, como adulteração de medidores e ligações clandestinas, o suficiente para abastecer por dez meses o município de Santo André, com 670000 habitantes, por exemplo (EXAME 2013).

Por outro lado, precisa-se fortalecer a cobertura das operadoras de serviços de telecomunicações necessárias para a transmissão de dados, a fim de cumprir todas as expectativas. Embora, para isso os temas inerentes ao processo de regulamentação, atualmente sob responsabilidade da Agência Nacional de Telecomunicações ou ANATEL, e da Agência Nacional de Energia Elétrica ou ANEEL, órgãos reguladores do setor elétrico, necessitem ser analisados.

A isto soma-se outros assuntos de importância que devem ser analisados para um bom desenvolvimento desta nova área. Alguns enfoques de trabalho futuros, na mesma linha de pesquisa do projeto realizado, são detalhados no último capítulo.

2.3 Resumo do Capítulo

As redes *Smart Grid* são apresentadas como uma forte tendência de modernização da atual rede elétrica. Esta nova rede chega para solucionar os problemas que a antiga rede elétrica vinha apresentando, especificamente na distribuição. Adicionando-se uma camada de comunicações, procura-se transmitir de forma eficiente os dados de controle e consumo de maneira bidirecional.

A substituição dos medidores de consumo elétrico, diminuirá os problemas de furto de energia e junto com os dispositivos de gestão de energia dentro da casa, trará a compreensão e conscientização do uso de energia, o que acarretará em um benefício tanto para o consumidor como para a companhia elétrica.

Assim como o uso da internet começou a ser de grande importância depois da introdução do computador, o mesmo acontecerá com o desenvolvimento de uma camada de comunicações que abrange uma série de equipamentos de medição e sensores de dados em tempo real, resultando em um significativo fluxo de dados que terá que ser processado todos os dias.

Portanto, estudos na análise de compressão de dados de consumo de potência fornecidos pelo medidor tornam-se muito importantes, assim como outros assuntos de cobertura, soluções de regulamentação, segurança, dentre outros.

Caracterização da Demanda Média Residencial e Teoria da Codificação de Fonte

Este capítulo caracteriza a curva de carga representativa de um consumidor residencial e a modelagem deste processo estocástico. Em seguida, define-se a base teórica das cadeias de Markov e seu desenvolvido para a aplicação na modelagem do perfil de atividades dos residentes. Na ultima seção, apresenta-se uma revisão geral da teoria de codificação de fonte, os tipos de compressão e técnicas pertencentes a estas classe.

3.1 Curva de Carga de um Consumidor Residencial

A curva diária de demanda é a representação de consumo da energia elétrica, ela expressa a variação da potência ou demanda no tempo ($Watts \times tempo$). Para sua representação, é preciso definir antes a demanda média em um período T , que é o valor da energia consumida E dividida pelo período, ou seja, E/T . Normalmente, as curvas de carga diárias dos consumidores são medidas em períodos de 15 à 30 *dias* em média, com registros diários de potência média em intervalos de 1 *hora*, ou 15 *minutos*. Na engenharia de distribuição de energia elétrica, trabalha-se normalmente com as demandas médias em intervalos de 15 *minutos* ou taxa de registro de 1/900 Hz, resultando em uma curva diária de 96 amostras ($n = \{1, 2, \dots, 96\}$)(Jardini 1999), como é apresentado na Fig. 3.1.

A curva apresentada na Fig. 3.1 ilustra um perfil típico de um consumidor residencial. Nesta curva é importante conhecer o pico (demanda máxima), o horário de pico e quanto cada carga contribui para o pico global (Jardini 1999).

Da curva, pode-se definir alguns parâmetros de interesse:

- **Demanda Máxima:**

A demanda máxima D_{max} é a maior demanda média na curva. Nos estudos de distribuição,

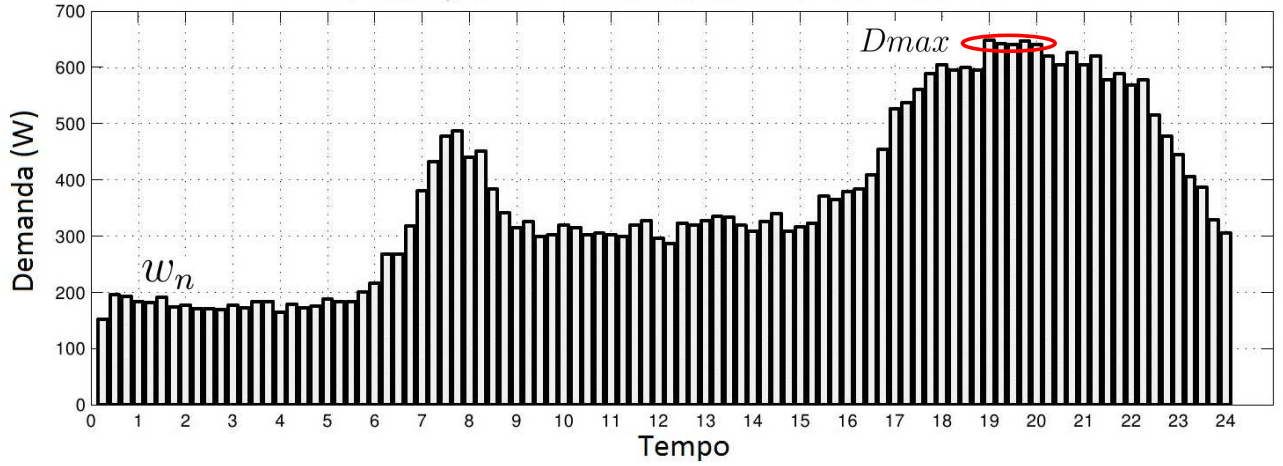


Figura 3.1: Curva de demanda média representativa do consumidor residencial ($Watts \times Tempo$), com registros diários em intervalos de $T = 15 \text{ minutos}$, gerando 96 amostras, e consumo de energia diária de 8,84 kWh/dia.

o instante de pico é definido ao redor das 20 : 00 horas com uma contribuição com valores de até 1013 W no pico global (Jardini 1999).

Como exemplo, na Fig. 3.1, o $Dmax = 650 \text{ W}$ encontra-se dentro do horário entre as 19 e 20 horas, respectivamente.

- **Energia consumida diária:**

A energia consumida diária E é a área da curva de demanda. Assumindo T em minutos, a energia é dada por:

$$E = \sum_n w_n \times \frac{T}{60} \text{ Wh} \quad (3.1)$$

Onde w_n é um valor de demanda no instante n .

O valor característico de energia consumida diária para um consumidor residencial é de 13,3 kWh (Jardini 1999). O valor para a Fig. 3.1 é de 8,84 kWh.

- **Potência média diária:**

Potência ou demanda média é obtida dividindo-se a energia diária por 24 horas. O valor característico de potência média diária para um consumidor residencial é de 0.555 kW (Jardini 1999).

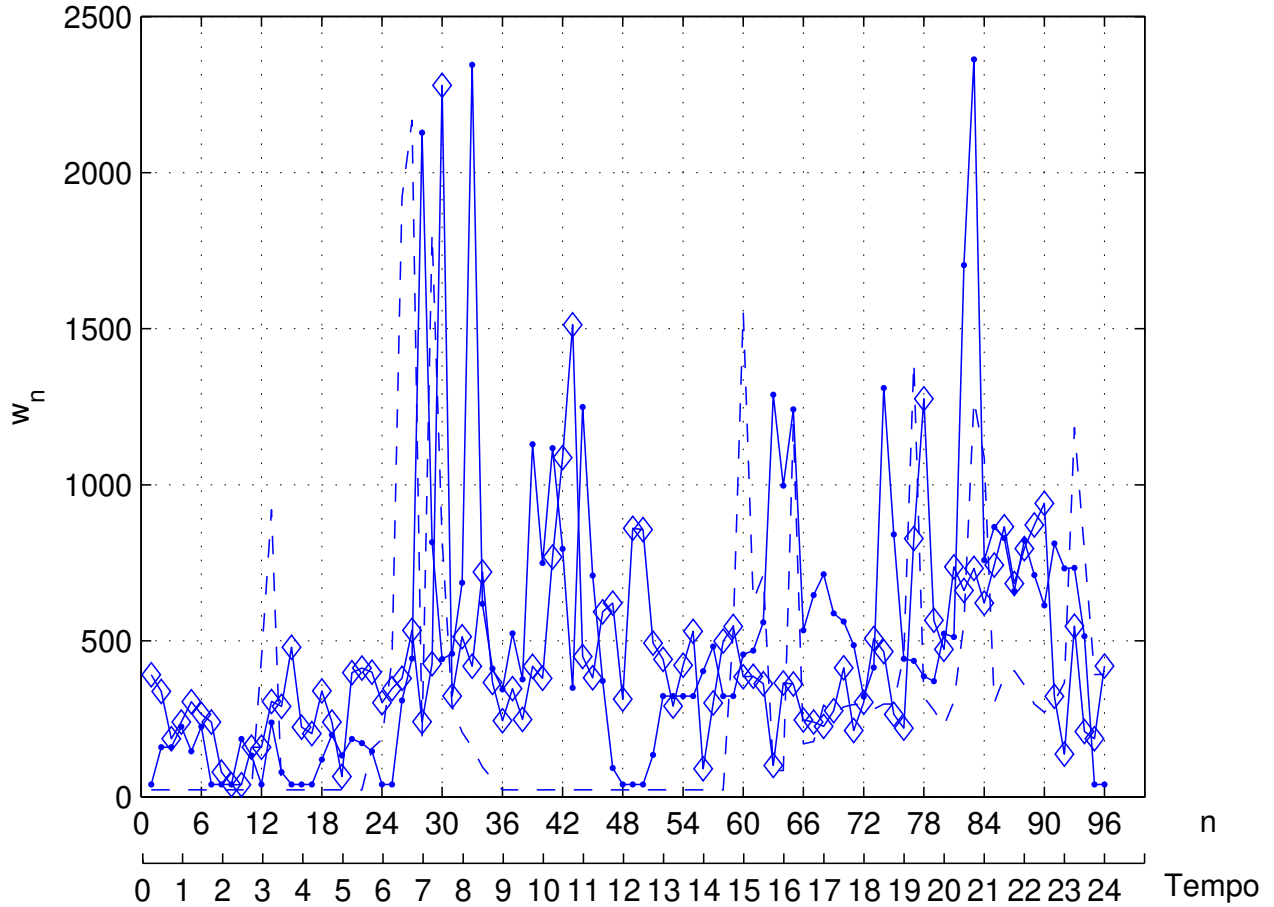


Figura 3.2: Saídas do processo aleatório de consumo elétrico. Cada realização representa um perfil individual de consumo residencial, com registros diários em intervalos de $T = 15$ minutos, gerando 96 amostras.

3.2 Modelagem de Fonte de Dados de Consumo de Energia

Entenda-se como fonte de dados, neste caso o medidor de consumo elétrico e os valores de demanda em potência média são os dados da fonte. Um conjunto de residências envia dados para um ponto central, os armazena em um disco, e os transmite diretamente pela rede até as empresas distribuidoras de energia elétrica como foi visualizado na Fig. 2.3.

Os dados de consumo podem ser representados formalmente como uma série de dados no tempo $D = \{D_n\}$, em que $D_n \in D$ é uma tupla (t_n, W_n) , no qual t_n representa o tempo e W_n o consumo (Wijaya et al. 2013). Dito de outra forma, a fonte de dados pode ser modelada matematicamente por um processo estocástico W_n , definido como sendo uma sequência infinita de variáveis aleatórias $(W_0, W_1, \dots)$ indexadas pelo tempo.

Este processo aleatório é discreto no tempo mas com valores contínuos ou DTCV (do inglês

"Discrete-Time/Continuous-Valued"). O intervalo entre as medidas é definido como T e iremos considerar n amostras desta sequência (Kay 2012).

No que refere ao nível de autocorrelação de um processo aleatório em tempo discreto, pode-se modelar a fonte da seguinte forma: (Covers & Thomas 2006, Gallager 1976):

- **Modelo de Ordem Zero:**

Neste modelo, cada medição é estatisticamente independente de todas as outras medições, e os possíveis valores da fonte ocorrem de forma equiprovável.

- **Modelo de Primeira Ordem:**

Neste caso, algumas medições ocorrem com mais frequência que outras, mas ainda são independentes.

- **Modelo de Segunda Ordem:**

Neste modelo, assume-se que existe alguma dependência entre as medições. Naturalmente, medições que estão próximas são mais dependentes do que aquelas que estão mais distantes.

Neste modelo, a medição W_n depende da medição W_{n-1} , mas é condicionalmente independente de todas as medições anteriores ($W_1, W_2, \dots, W_{n-2}$).

- **Modelo de Terceira Ordem:**

Este é uma extensão do modelo anterior. A presente medição W_n depende das duas medições anteriores (W_{n-2}, W_{n-1}), mas é condicionalmente independente de todas as outras medições: ($W_1, W_2, \dots, W_{n-3}$).

- **Modelo Geral:**

As propriedades estatísticas deste modelo são muito complexas para serem consideradas práticas. Este modelo é interessante apenas do ponto de vista teórico.

3.3 Modelo de Markov e Membros Ativos de uma Residência

Em uma cadeia de Markov, X_n é chamado de estado no tempo n , e pode assumir qualquer valor do espaço de estados $\mathcal{X}$. Este processo é caracterizado por seu estado inicial e uma matriz de transição $\mathbf{P} = [P_{x_n x_{n+1}}]$, em que $x_n, x_{n+1} \in \mathbb{N}$, e $P_{x_n x_{n+1}} = Pr \{X_{n+1} = x_{n+1} | X_n = x_n\}$

(Covers & Thomas 2006), representado por meio de um grafo dirigido, onde cada aresta é rotulada com as probabilidades de transição de um estado para outro, sendo esses estados os nós conectados pelas arestas, como é ilustrado na Fig. 3.3.

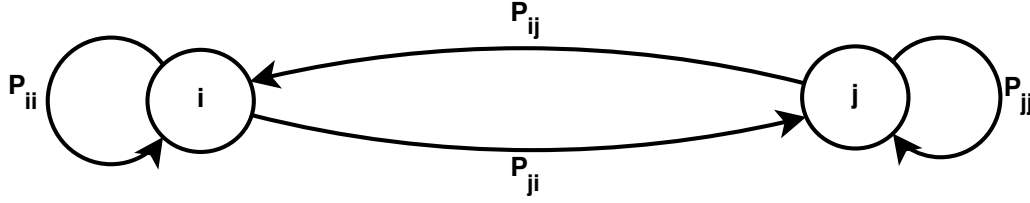


Figura 3.3: Diagrama de estados de uma cadeia de Markov com dois estados ou $m = 1$. Sendo $x_n = i$ e $x_{n+1} = j$, P_{ij} e P_{ji} são as probabilidades de transição de sair de um estado para outro, P_{ii} e P_{jj} são as probabilidades de manter-se no mesmo estado.

Em uma cadeia de Markov de primeira ordem (Covers & Thomas 2006, Kay 2012), a seguinte relação é válida:

$$Pr(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_1 = x_1) = Pr(X_{n+1} = x_{n+1} | X_n = x_n) \quad (3.2)$$

para todo $x_1, x_2, \dots, x_n, x_{n+1} \in \mathcal{X}$.

Por outro lado, em uma cadeia de Markov de ordem m , na qual o estado futuro depende de m estados passados, a seguinte relação é válida

$$\begin{aligned} Pr(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_1 = x_1) = \\ Pr(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_{n-m} = x_{n-m}) \end{aligned} \quad (3.3)$$

para $n > m$.

Além disso, uma cadeia de Markov é dita ser invariante no tempo (homogênea ou estacionária) se a probabilidade condicional $p(x_{n+1}|x_n)$ não depende de n , definida como:

$$Pr\{X_{n+1} = x_{n+1} | X_n = x_n\} = Pr\{X_2 = x_{n+1} | X_1 = x_n\} \quad (3.4)$$

No caso de depender de n , tal cadeia de Markov é dita ser não homogênea, tendo uma matriz de transição diferente para cada instante de tempo, definida como:

$$Pr\{X_{n+1} = x_{n+1} | X_n = x_n\} \neq Pr\{X_2 = x_{n+1} | X_1 = x_n\} \quad (3.5)$$

Desta forma, a cadeia de Markov tem muitas aplicações em modelos estatísticos de processos do mundo real, sendo amplamente utilizada para modelar sequências de eventos. Um exemplo claro é a geração de dados de ocupação dos membros de uma família, apresentada em (Richardson et al. 2008). Os dados foram obtidos por meio de entrevistas sobre como as pessoas utilizam seu tempo, chamado TUS (do inglês *Time-Use Survey*). Um exemplo do registro TUS é apresentado na Fig. 3.4 ¹.

Nota-se que existe uma atividade mínima na madrugada, algumas atividades durante o dia e mais atividades durante a noite (Richardson et al. 2008).

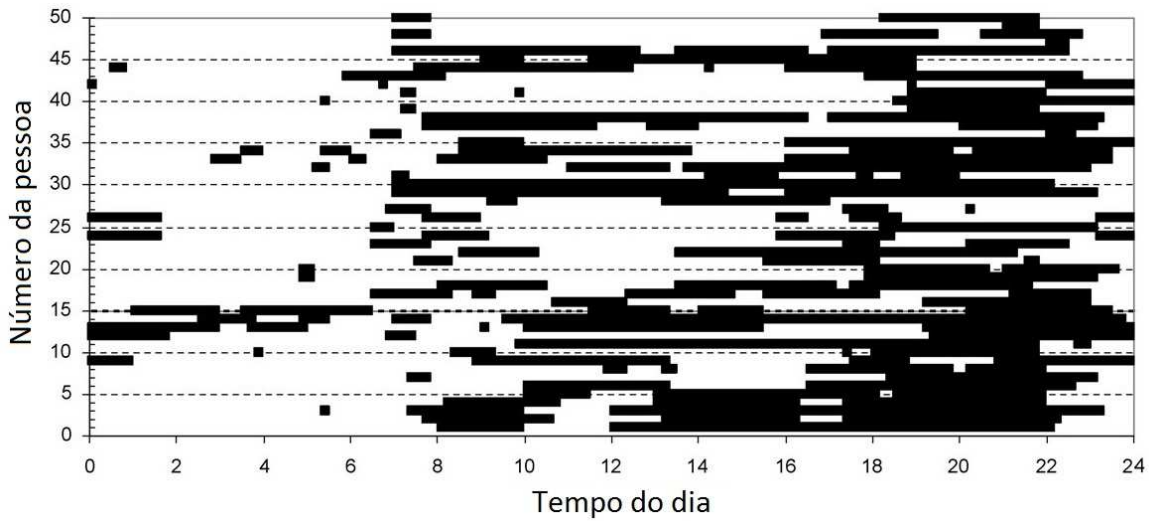


Figura 3.4: Cinquenta exemplos de perfis de ocupação ativos, de uma família de 1 membro em um dia da semana, extraídos a partir dos dados TUS

A partir destes registros de dados TUS, é possível determinar os parâmetros das probabilidades de transição de um processo Markoviano de primeira ordem $m = 1$ (ou fonte com modelo estatístico de segunda ordem) examinando as probabilidades condicionais para cada intervalo de tempo, gerando uma cadeia de Markov não homogênea. Os estados deste processo são o número de residentes ativos. Para os dados da Fig. 3.4, temos até 6 membros como número máximo (Richardson et al. 2008).

Todas as transições entre as amostras n e $n + 1$ são examinadas e o número total $num_{ij}(n)$ de transições entre o estado i e j é determinado. O número total de mudanças a partir do estado i é dado por $num_i(n) = \sum_{j=0}^f num_{ij}(n)$. Assim, para uma família de f membros, as probabilidades $P_{ij}(n)$ para cada instante de tempo são armazenadas nas matrizes de probabilidades de transição $\mathbf{P}$ (Widen & Wackelgard 2010). Cada elemento desta matriz pode ser escrito da seguinte forma:

¹fonte: (Richardson et al. 2008)

$$P_{ij}(n) = \frac{num_{ij}(n)}{num_i(n)} \quad (3.6)$$

Como exemplo de descrição, assume-se uma residência de $f = 5$ membros, o estado inicial para esta família será definida pela distribuição inicial apresentada na tabela 3.1 ². Nota-se que é muito provável que ninguém permaneça ativo no primeiro período do dia.

Tabela 3.1: Distribuição inicial de ocupação ativa para uma família de 5 membros. Distribuição resultante dos registros TUS.

Membro ativo	0	1	2	3	4	5
$\Pr(X_0 = \text{Membro ativo})$	0,680	0,176	0,120	0,008	0,008	0,008

Os seguintes estados estarão definidos por uma matriz de probabilidade de transição de Markov não homogênea, ou seja, uma para cada instante de tempo n . Seguindo o exemplo com a mesma família de 5 integrantes, as respectivas matrizes P_{ij} com $i, j \in \{0, 1, 2, 3, 4, 5\}$ são definidas e ordenadas na tabela 3.2 ³ para 3 diferentes instantes do dia, especificamente às 3 : 00, 12 : 00 e 19 : 00 horas respectivamente, por facilidade de apresentação. Nota-se que para estes instantes de tempo, as matrizes de transição são diferentes.

Assumindo que o estado i seja de 2 residentes ativos às 12 : 00 horas, é muito provável que se mantenham as mesmas 2 pessoas para o seguinte período de tempo, e bem pouco provável que chegue mais um ou fique um residente ativo na casa.

Desta maneira, um perfil aleatório sobre quantos membros de uma residência estão ativos durante o dia é gerado e apresentado na Fig. 3.5.

A Fig 3.6 apresenta o total de membros ativos. Desta curva, pode-se notar o mesmo comportamento da curva de carga de demanda média residencial. Estes dados são influenciados pelas atividades dos membros da casa.

3.4 Agregação Temporal

A agregação temporal representa uma segmentação vertical na série de tempo a fim de reduzir a numerosidade dos dados.

²(Richardson et al. 2008)

³(Richardson et al. 2008)

Tabela 3.2: Matriz probabilidade de transição não homogênea de i à j ocupantes ativos $P_{ij}(n)$, para às 03 : 00, 12 : 00 e 19 : 00 horas do dia respectivamente, em uma residência de 5 membros. Matrizes resultantes dos registros TUS.

03:00						
i,j	0	1	2	3	4	5
0	0,992	0,008	0,000	0,000	0,000	0,000
1	0,286	0,714	0,000	0,000	0,000	0,000
2	0,000	0,000	0,000	0,000	0,000	0,000
3	0,000	0,000	0,000	0,000	0,000	0,000
4	0,000	0,000	0,000	0,000	0,000	0,000
5	0,000	0,000	0,0000	0,000	0,000	0,000
12:00						
i,j	0	1	2	3	4	5
0	0,914	0,086	0,000	0,000	0,000	0,000
1	0,067	0,933	0,000	0,000	0,000	0,000
2	0,000	0,105	0,737	0,158	0,000	0,000
3	0,000	0,100	0,100	0,800	0,000	0,000
4	0,000	0,000	0,000	0,286	0,714	0,000
5	0,000	0,000	0,0000	0,000	0,000	1
19:00						
i,j	0	1	2	3	4	5
0	1	0,000	0,000	0,000	0,000	0,000
1	0,000	0,667	0,250	0,083	0,000	0,000
2	0,000	0,103	0,724	0,172	0,000	0,000
3	0,000	0,000	0,269	0,577	0,154	0,000
4	0,057	0,000	0,029	0,343	0,514	0,057
5	0,000	0,000	0,0000	0,176	0,294	0,529

Em consequência da agregação temporal, uma nova versão de série temporal $\dot{D} = \{\dot{D}_1, \dot{D}_2, ..\}$ é gerada da original D , cujas duplas $(\dot{t}_n, \dot{W}_n)$, serão dadas por (Wijaya et al. 2013):

$$\dot{t}_n = t_{n \times T} \quad (3.7)$$

$$\dot{W}_n = \frac{1}{T} \sum_{(n-1) \times T + 1}^{n \times T} W_n \quad (3.8)$$

No entanto, estes novos dados $\dot{D}$ continuam sendo caracterizados como um processo DTCV, no qual a nova série $\dot{W}_n$ assume valores do alfabeto $\mathcal{W}$.

Este processo reproduz o funcionamento do registrador de cada medidor de consumo de

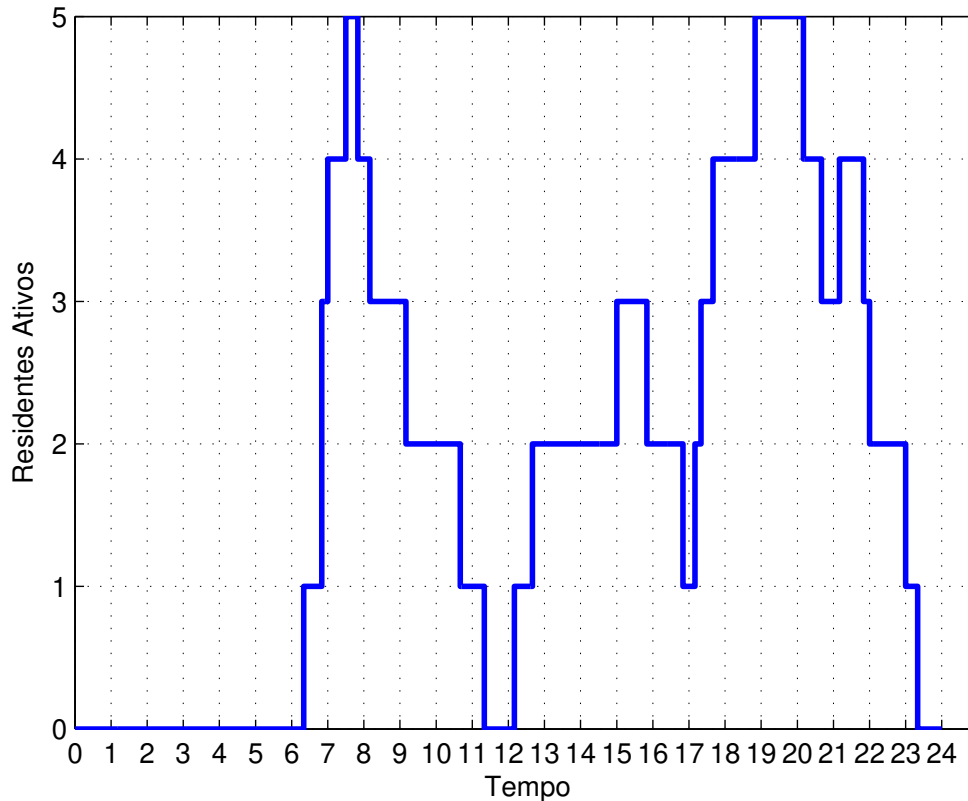


Figura 3.5: Perfil de ocupação dos membros ativos em uma residência de 5 membros. Dados modelados usando matrizes de transição de cadeia de Markov não homogênea, resultantes dos registros TUS.

energia elétrica, que acumula os pulsos num dado intervalo de tempo especificado, determinando a energia e , por consequência, a potência média no intervalo. Esse valor de potência é armazenado na memória do dispositivo, seguindo tabelas horárias, diárias e mensais. Um medidor típico pode armazenar dados referentes a 38 *dias* de medição com potências definidas em intervalos de 15 *minutos* (Jardini 1999).

3.5 Teoria de Codificação da Fonte de Dados

Em 1948, foi formulado por Claude E. Shannon a teoria de compressão de dados da fonte. A teoria abrange duas classes de conceitos de compressão: as sem perdas e com perdas.

A teoria de compressão de dados sem perdas estabelece os limites fundamentais sobre o desempenho de todos os algoritmos de compressão de dados. Shannon estabeleceu que há um limite fundamental à compressão de dados sem perdas. Esse limite é nomeado de taxa de entropia, denotado por $\mathcal{H}$. O valor exato de $\mathcal{H}$ depende da informação da fonte, mais especificamente da sua natureza estatística. Assim, é possível comprimir a fonte, sem perdas, com taxa de compressão perto de $\mathcal{H}$. Ou seja, a melhor taxa de compressão de dados sem perdas é a taxa

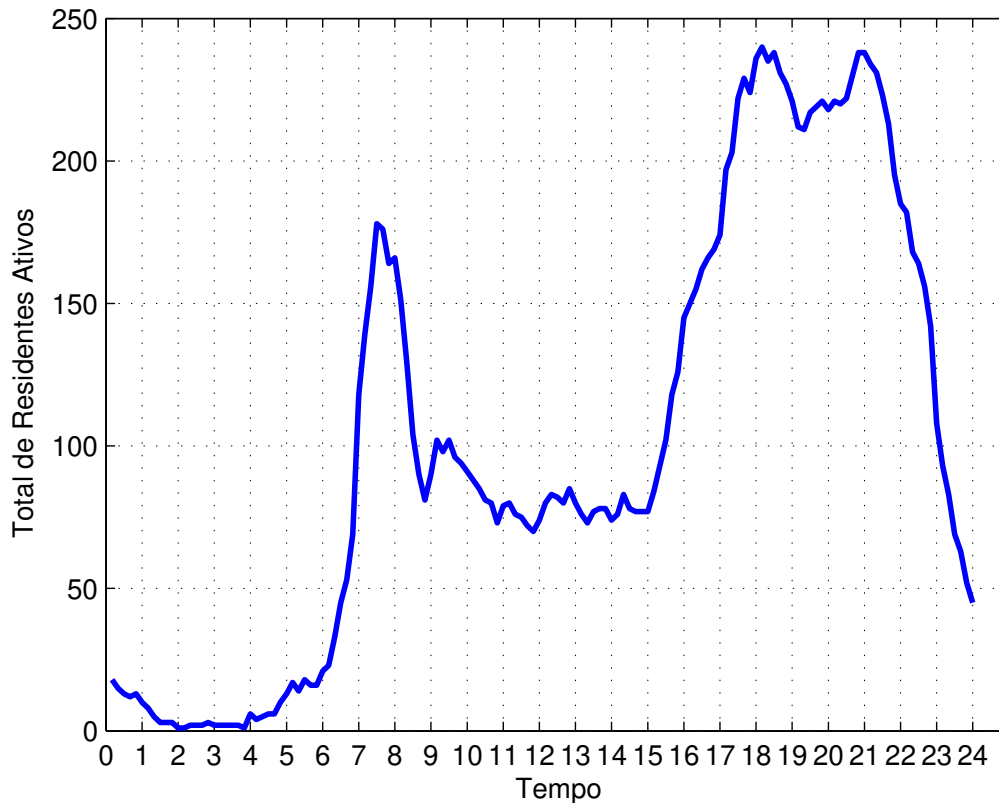


Figura 3.6: Curva total de residentes ativos de 123 residências, e famílias de 1 à 6 membros. Correspondência entre o comportamento das atividades dos membros da casa com a curva de demanda elétrica residencial.

da entropia (Covers & Thomas 2006).

A teoria de compressão de dados com perdas é conhecida como a teoria da taxa de distorção. Na compressão de dados com perdas, os dados recuperados não são os mesmos do que os originais e o desempenho pode ser medido pelo compromisso entre a compressão e distorção (Covers & Thomas 2006).

3.5.1 Compressão em Amplitude - Quantização

Como foi definido, em muitas aplicações tais como processamento digital de sinais analógicos, onde o alfabeto da fonte não é discreto, o número de bits necessários para a representação de cada saída da fonte não é finito.

A fim de processar a saída da fonte digitalmente, a fonte tem que ser quantizada em um número finito de níveis. Este processo introduz alguma distorção inevitavelmente. No entanto, quando a distorção é introduzida de forma controlada no sinal sem comprometer o seu uso ou aplicação, é possível alcançar taxas de compressão mais elevadas do que as abordagens de

compressão sem perdas. Ou seja, a degradação do sinal pode ser tal que o sinal ainda é útil (Proakis & Salehi 1997, Tcheou, Lovisolo, Ribeiro, Da Silva, Rodrigues, Romano & Diniz 2014).

Técnicas gerais de conversão A/D, tais como: modulação por codificação de pulso ou PCM (do inglês "Pulse Code Modulation"), modulação por codificação de pulso diferencial DPCM (do inglês "Differential Pulse Code Modulation", modulação delta, quantização uniforme, quantização não uniforme, e quantização vetorial pertencem à esta classe. Nesta compressão, o desempenho pode ser medido pelo compromisso entre o ganho de compressão e distorção (Proakis & Salehi 1997, Tcheou et al. 2014).

Em seguida, duas técnicas gerais de quantização serão apresentadas (Proakis & Salehi 1997):

- **PCM**

Este método é o mais utilizado para representar digitalmente sinais analógicos amostrados. É um formato padrão para aplicações de áudio digital, computadores, discos compactos e áudio digital.

Em um fluxo PCM, a amplitude do sinal analógica $\dot{W}_n$ é amostrada regularmente em intervalos uniformes, e cada amostra é quantizada ao valor mais próximo dentro de um intervalo. Ou seja, este é dividido dentro de N regiões $\mathcal{R}_i$ não sobrepostas, com $1 \leq i \leq N$ chamado intervalos de quantização, e dentro de cada região um ponto chamado nível de quantização é escolhido. Estes níveis de quantização definem o alfabeto finito $\dot{\mathcal{W}} = \{\dot{w}_i\} = \{\dot{w}_1, \dot{w}_2, \dots, \dot{w}_N\}$ em cada n . Então todos os valores da variável aleatória $\dot{W}_n$ que caem dentro da região $\mathcal{R}_i$ são quantizados no i -ésimo nível de quantização, denotado por $\dot{W}_{in}$ (Proakis & Salehi 1997, Tcheou et al. 2014), dado por:

$$\dot{W}_n \in \mathcal{R}_i \Leftrightarrow Q(\dot{W}_n) = \dot{W}_{in}, \quad (3.9)$$

em que $\dot{W}_{in} \in \dot{\mathcal{W}}$.

Quando os valores máximos e mínimos do sinal de entrada são conhecidos, pode-se projetar um quantizador cuja região abranja todo o domínio do sinal. A Fig. 3.7 ilustra um exemplo de PCM.

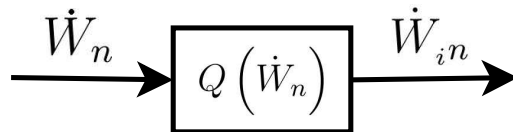


Figura 3.7: Esquema PCM

• DPCM

Neste esquema, um algoritmo faz a predição da amostra seguinte baseada na amostra anterior, e quantiza usando PCM, somente a diferença entre esta predição e o valor atual. Se esta predição é razoável, muito poucos bits serão necessários para representar a informação (Tcheou et al. 2014).

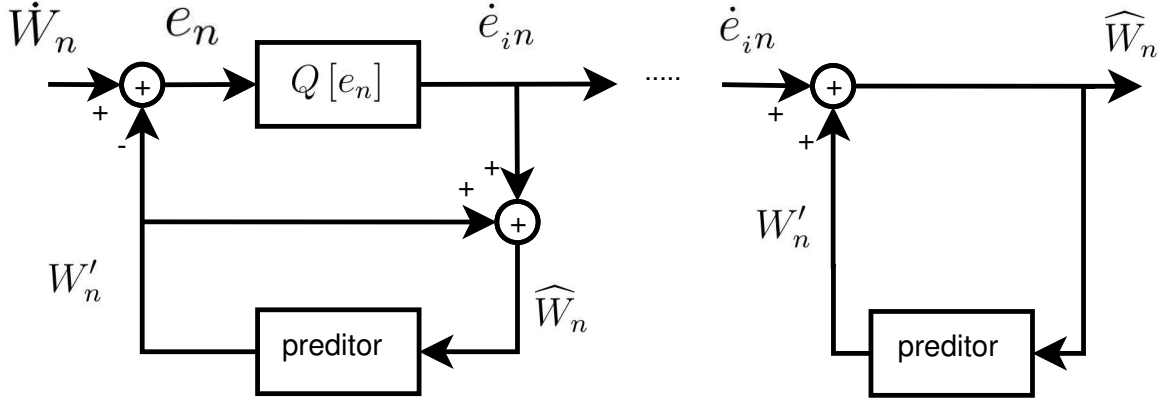


Figura 3.8: Esquema DPCM

O esquema pode ser definido matematicamente como segue:

$$\begin{aligned} e_n &= \dot{W}_n - W'_n \\ \dot{e}_{in} &= Q[e_n] \\ \widehat{W}_n &= W'_n + \dot{e}_{in} \end{aligned} \quad (3.10)$$

A distorção é dada por:

$$\mathcal{D} = E \left[\left(\dot{W}_n - Q(\dot{W}_n) \right)^2 \right], \quad (3.11)$$

em que $E[.]$ é operador esperança.

Em consequência, a relação sinal-ruído de quantização ou SQNR (do inglês "Signal-to-Quantization-Noise Ratio") (Proakis & Salehi 1997) é medida como:

$$SQNR|_{dB} = 10 \log_{10} \frac{E[\dot{W}_n^2]}{\mathcal{D}} \quad (3.12)$$

É obvio que, com um número grande de níveis de quantização, será esperado um erro de quantização muito pequeno (menor distorção, maior SQNR).

Compressão dos dados digitais - Codificação

Os níveis quantizados são codificados utilizando b bits para cada nível de quantização. Se N é uma potência de 2 ou $N = 2^b$, então b bits são requeridos para a representação de cada nível.

O teorema de codificação sem perdas estabelece que para uma reconstrução perfeita da fonte é necessário usar um código com uma taxa maior do que a entropia da fonte. Vários são os algoritmos de codificação sem perdas, o mais conhecido deles é o codificador de Huffman.

• Codificação Huffman

A codificação de Huffman é um método de codificação estatística que se baseia na frequência de ocorrência de cada símbolo. Os padrões que ocorrem com menor frequência serão codificados com mais bits, já os de maior frequência serão codificados com menos bits.

O comprimento médio da palavra código $\mathcal{L}$ para uma codificação de Huffman satisfaz a seguinte inequação:

$$\mathcal{H} \leq \mathcal{L} < \mathcal{H} + 1 \quad (3.13)$$

com $\mathcal{L}$, definida como:

$$\mathcal{L} = \sum_{i=1}^N l_i p(\dot{w}_{in}), \quad (3.14)$$

na qual l_i é o comprimento do i -ésimo nível de quantização, e $p(\dot{w}_{in})$ é a probabilidade de ocorrência do i -ésimo nível de quantização.

• Taxa de Entropia de uma Fonte

Para uma fonte de informação que contém uma sequência de variáveis aleatórias $\dot{W}_{in}$ oscilando sobre um alfabeto finito $\dot{w}_{in} \in \dot{\mathcal{W}}$, e função de distribuição de primeira ordem $p(\dot{w}_{in}) = Pr(\dot{W}_{in} = \dot{w}_{in})$ com $i = \{1, 2, \dots, N\}$, a entropia de primeira ordem é dada por:

$$\mathcal{H}_1(\dot{W}_{in}) = - \sum_{\dot{w}_{in} \in \dot{\mathcal{W}}} p(\dot{w}_{in}) \log_2 \left(\frac{1}{p(\dot{w}_{in})} \right) \quad (3.15)$$

Quando todos os níveis de quantização são igualmente prováveis, ou seja, $p(\dot{w}_{in}) = 1/N$, em todos os instantes de tempo, a taxa de entropia é dada por $\log_2 N$.

Para processos aleatórios, que podem ser representados por uma cadeia de Markov de primeira ordem ($m = 1$), ou cuja fonte pode ser modelada como de segunda ordem, a taxa

de entropia desta fonte é definida como a entropia condicional $\mathcal{H}_2(\dot{W}_n|\dot{W}_{n-1})$, dada por:

$$\mathcal{H}_2(\dot{W}_n|\dot{W}_{n-1}) = - \sum_{\dot{w}_{n-1} \in \dot{\mathcal{W}}} p(\dot{w}_{n-1}) \sum_{\dot{w}_n \in \dot{\mathcal{W}}} p(\dot{w}_n|\dot{w}_{n-1}) \log_2 p(\dot{w}_n|\dot{w}_{n-1}) \quad (3.16)$$

em que, $\dot{W}_n \sim Pr(\dot{W}_n)$ e $\dot{W}_{n-1} \sim Pr(\dot{W}_{n-1})$ são duas variáveis aleatórias com suas respectivas funções de distribuição, e $p(\dot{w}_n|\dot{w}_{n-1})$ é a probabilidade condicional de que a presente amostra seja $\dot{w}_n$ dada que a amostra anterior foi $\dot{w}_{n-1}$.

De maneira similar, se o processo aleatório é descrito pelas cadeias de Markov de segunda ordem ($m = 2$), ou cuja fonte de dados pode ser modelada de terceira ordem, a taxa de entropia é definida como a entropia condicional $\mathcal{H}_3(\dot{W}_n|\dot{W}_{n-1}, \dot{W}_{n-2})$, dada por:

$$\begin{aligned} \mathcal{H}_3(\dot{W}_n|\dot{W}_{n-1}, \dot{W}_{n-2}) = & - \sum_{\dot{w}_{n-2} \in \dot{\mathcal{W}}} p(\dot{w}_{n-2}) \sum_{\dot{w}_{n-1} \in \dot{\mathcal{W}}} p(\dot{w}_{n-1}|\dot{w}_{n-2}) \\ & \sum_{\dot{w}_n \in \dot{\mathcal{W}}} \log_2 p(\dot{w}_n|\dot{w}_{n-1}, \dot{w}_{n-2}) \end{aligned} \quad (3.17)$$

3.6 Resumo do Capítulo

Neste capítulo foram apresentados alguns parâmetros ou fatores característicos da curva de carga tais como a demanda máxima e a energia diária consumida.

O conceito de demanda elétrica foi estendido e definido formalmente como um processo aleatório, cujas realizações representam os perfis individuais de consumo diário residencial, que são úteis para calcular a distribuição de demanda total de uma região.

Seguindo esta definição foi visto que a demanda pode ser caracterizada como um processo de tempo discreto com valores contínuos, ou DTCV, uma vez que é definida em intervalos de tempo estabelecidos, tipicamente de 15 *minutos*, gerando uma sequência de dados de medições com valores contínuos ou de alfabeto não finito.

Sendo definidos os dados da fonte e dada a importância do estudo da compressão de dados sobre os registros de consumo elétrico, foi descrito a base da teoria de codificação de fonte. Assim, foi visto que o processo de compressão é baseado em dois passos: na transformação

analógica/digital para a compressão na amplitude utilizando os esquemas PCM e DPCM, e a compressão dos dados digitais usando técnicas de codificação como a de Huffman.

Finalmente, dado que a teoria especifica o limite mínimo de compressão teórica definida pela entropia. A entropia de um processo foi definida para um processo Markoviano.

Limite Máximo de Compressão de Dados do Consumo Elétrico

Neste capítulo, apresenta-se a descrição dos simuladores utilizados para a modelagem das curvas de demandas e a avaliação destes com dados reais.

4.1 Modelo de Compressão de Dados de Consumo Elétrico

Um modelo de compressão de dados é apresentado no diagrama de bloco da Fig 4.1. O esquema apresenta o processamento dos dados de consumo elétrico baseado em uma representação simbólica. Os dados de consumo são representado como uma serie temporal D . Em seguida, os dados entram por um processo de segmentação vertical, no qual a representação da saída será um equivalente sintetizado no tempo representado por $\hat{D}$ para períodos de registro de dados T .

Os dados passam por um processo de segmentação horizontal. Neste ponto, muda-se a granularidade dos valores de consumo a fim de processar a saída da fonte de forma digital, quantizando o consumo $\dot{W}_n$ a um número finito de níveis $\dot{W}_{in}$, e que por sua vez são codificados a fim de representá-los com a menor quantidade de bits possíveis.

4.2 Modelagem de Demanda de Energia Elétrica Residencial

Dado que o conhecimento da demanda de energia é de vital importância para o projeto de aparelhos inteligentes, dimensionamento da rede de distribuição, modelagem da resposta de demanda, alimentação inteligente de veículos elétricos, gerenciamento da energia da casa, previsão dos tipos de tecnologias que farão parte do sistema, e evolução do comportamento na utilização de energia residencial, vários projetos de campo têm sido implementados para adquirir dados

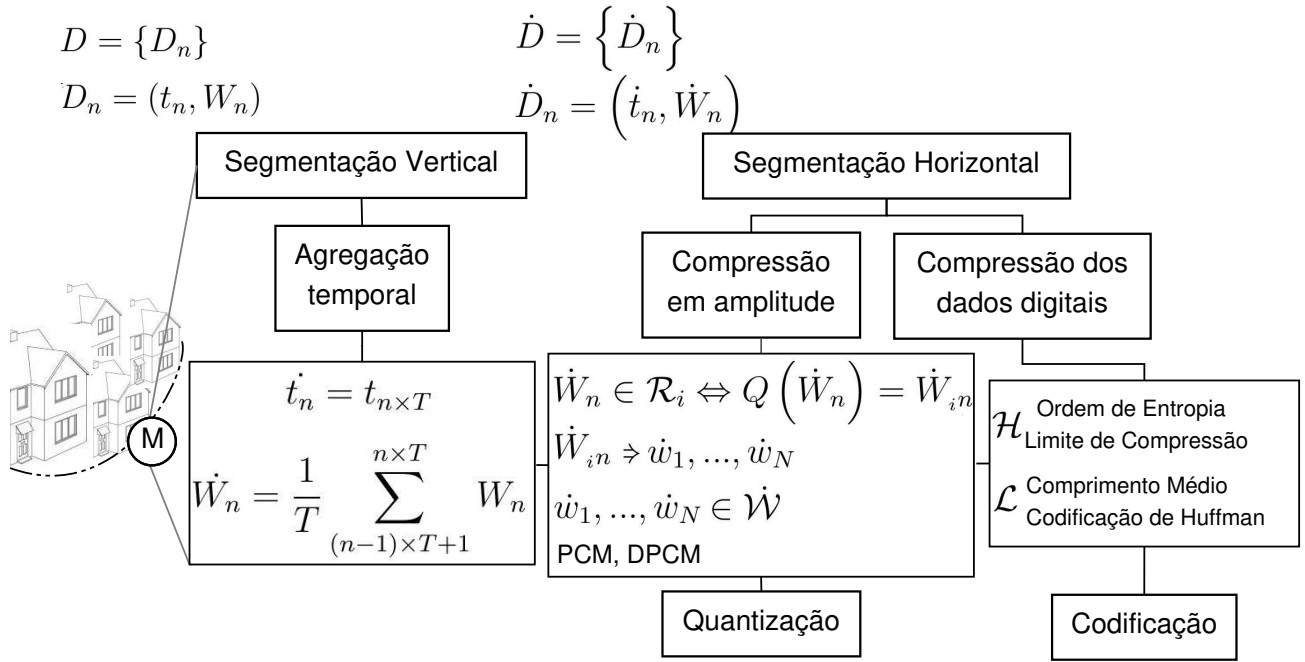


Figura 4.1: Esquema de compressão dos dados de consumo elétrico residencial.

de consumo (Kolter & Johnson 2011, Ardakanian et al. 2011, Energymyndigheten 2013).

A modelagem do consumo de eletricidade deve ser o mais simples possível. Muitos fatores determinam a carga da casa, tais como: horário do dia; meio ambiente e fatores geográficos; o tipo de eletrodoméstico e seu perfil de uso; o número de ocupantes da casa e seu perfil de atividades; e até as dimensões da casa. Um modelo simples pode necessariamente ignorar um ou mais fatores.

Para isso, dados relacionados a indicadores de consumo, potências e distribuições estatísticas dos eletrodomésticos, que são resultados de campanhas de pesquisa são necessários.

Duas modelagens de curva de carga foram utilizadas, estas são descritas a seguir:

- O primeiro modelo baseia-se nas informações de indicadores de parâmetros elétricos e estatísticos como a posse de eletrodomésticos e hábitos de uso dos aparelhos elétricos de consumidores residenciais do Brasil, fornecida pela PROCELINFO ("Centro Brasileiro de Informação de Eficiência Energética"), o qual inclui uma ferramenta de visualização das curvas resultantes dessas estatísticas, chamado SINPHA ("Sistema de Informação de Posse e Hábitos de Uso de Aparelhos Elétricos").

De forma geral, a confecção do perfil de carga, utiliza a potência média, duração média e hora pico média de utilização de tais aparelhos. Para a modelagem da curva, a aleatoriedade das cargas segue uma distribuição Gaussiana. A metodologia baseia-se em (Capasso, Grattieri, Lamedica & Prudenzi 1994, PROCEL & ELETROBRÁS 2005, EFLUL 2013, PROCELINFO 2013, Nunes et al. 2010, Jardim 1999).

- O segundo modelo é mais complexo e detalhado, levando em consideração os estados de ocupação dos consumidores definida por um modelo Markoviano, resultante de estatísticas de uso do tempo ou TUS.

A fim de validar o simulador, foi utilizado o conjunto de medições de consumo de baixa tensão feitas em campo, fornecido para fins de estudo na área de desagregação de energia, nomeado REED (do inglês "*Reference Energy Disaggregation Data Set*"), e disponibilizados em (Kolter & Johnson 2011). Estes conjuntos de dados registraram medições para 6 casas durante 3 meses aproximadamente a uma taxa de registro de 1 Hz.

Embora estas medições reais sejam de alta resolução, em certos momentos do dia não apresentaram medições de dados de forma contínua, ou seja, mostram intervalos de tempo sem medições. Além disso, o indicador de tempo de uma determinada medição é fornecida em formato UTC (do inglês "*Universal Time Coordinated*"), requerendo um processo inicial de conversão, a fim de identificar as potências correspondentes em um determinado instante do dia. Tudo isso, dificultou a escolha dos melhores e mais completos perfis. Assim, para manter a mesma referência, os simuladores foram feitos para a mesma quantidade de residências.

De forma geral, para o projeto de demanda de energia elétrica é importante modelar cada um de seus componentes como: o perfil de demanda de cada residente D_{R_r} , $r \in \mathbb{N}$; e seus equipamentos individuais D_{A_a} , $a \in \mathbb{N}$. O consumo de energia de cada residência é determinado pela energia elétrica consumida através do tempo de cada aparelho que esteve em uso.

O uso destes aparelhos pode ter diferentes finalidades, como: aquecimento de água (chuveiro), preparação e conservação de alimentos (liquidificador, micro-ondas, batedeira, refrigerador, congelador), condicionamento ambiental (ventilador, circulador e condicionador de ar), iluminação (lâmpada), lazer (equipamento de som, rádio elétrico, vídeo cassete, dvd, vídeo game, tv), higiene e limpeza (lava roupa, ferro de passar), trabalho (computador, impressora), etc.

O uso de eletricidade de uma determinada residência surge a partir da agregação de um conjunto finito de cargas "*on-off*" dada por 4.1. Em consequência, a demanda total D_{m_n} , será a demanda média entre todas as residências com seus respectivos aparelhos, dado por 4.2.

$$D_{R_r} = \sum_{a \in \mathbb{N}} D_{A_a} \quad (4.1)$$

$$D_{m_n} = \frac{1}{|r|} \sum_{r \in \mathbb{N}} D_{R_r} \quad (4.2)$$

4.2.1 Distribuição Gaussiana na Aleatoriedade do Uso da Carga

Este modelo foi o primeiro a ser idealizado e leva em consideração os indicadores resumidos na tabela 4.1.

Tabela 4.1: Especificações de aparelhos domésticos residenciais - Modelo 1

Aparelhos	Potência Média (W)	Hora Pico Média (h)	Duração Média (min)
Geladeira	90	-	-
Chuveiro	3500	7, 20	8
Micro-ondas	1200	9, 12, 19	10
Ferro	1000	10, 17	10
Máquina de lavar	500	9, 17	30
Computador	180	10, 18	300
Ventilador	120	12	240
Televisor	110	11, 21	180
DVD	100	22	60
Lâmpada	100	20	240
Aparelho de som	45	10, 18	240

O processo de geração da curva de demanda é detalhado no diagrama da Fig. 4.2. Inicialmente, sorteia-se a hora que o usuário inicia a utilização do aparelho seguindo uma distribuição normal, em seguida, as colunas de potência média e duração média são utilizadas para designar a demanda de cada aparelho durante um intervalo de tempo no dia.

Cada carga segue uma distribuição normal que tem como média o valor da coluna de hora pico e um desvio padrão arbitrário (2 horas), garantindo que pelo menos 60% dos sorteios fiquem em torno da hora pico.

Os dados resultantes foram obtidos para diferentes períodos de registro de consumo, especificamente para $T = \{1\text{seg}, 1\text{ min}, 5\text{ min}, 10\text{ min}, 15\text{ min}\}$ a fim de analisar o desempenho de compressão alcançada com taxas diferentes à taxa típica de 15 min. Como exemplo, considerando demandas médias de 15 minutos e supondo que um chuveiro com potência de 3,5 kW é ligado por 8 minutos, tem-se uma potência média em 15 minutos de $3,5 \cdot 8/15 = 1,86\text{ kW}$.

Com o propósito de exibir a aleatoriedade das curvas de cargas resultantes, a Fig 4.4, apresenta 10 perfis de cargas individuais para $T = 1\text{ minuto}$.

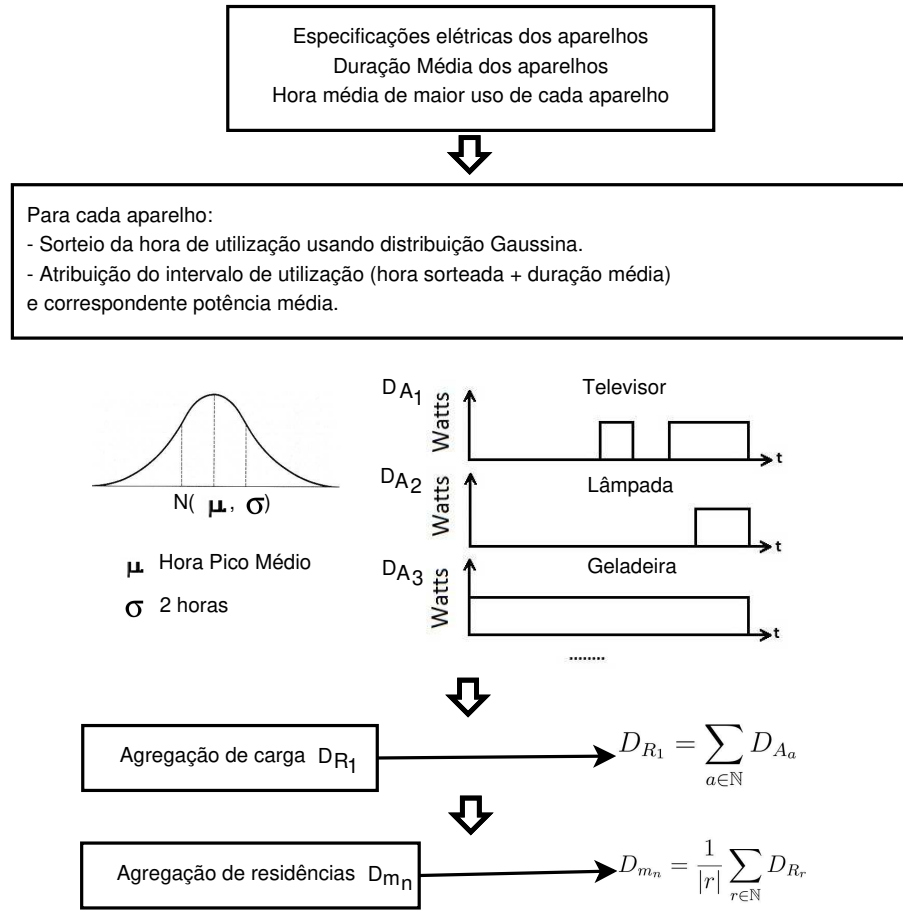


Figura 4.2: Esquema do processo de modelagem de demanda considerando uma distribuição normal para a aleatoriedade de cada aparelho.

4.2.2 Modelo Markoviano na Geração de Atividades dos Residentes

Este modelo utiliza as seguintes informações: posse de eletrodomésticos, fatores de potência média, duração média de uso, número de membros da casa, perfil de ocupação, distribuição de probabilidade de uso de cada aparelho durante o dia e seu modo de operação. Estes indicadores são resumidos na tabela 4.2.

O processo de geração da curva de demanda detalhado no diagrama da Fig. 4.3 é feito em duas partes, como segue:

- Inicia-se com a criação do perfil de estado de ocupação dos membros da casa.
- Nesta etapa, considera-se os eletrodomésticos usados pelos membros da casa. Estes são sorteados segundo a distribuição da posse de eletrodomésticos encontrado na tabela 4.2, e são classificados pelo modo de operação:

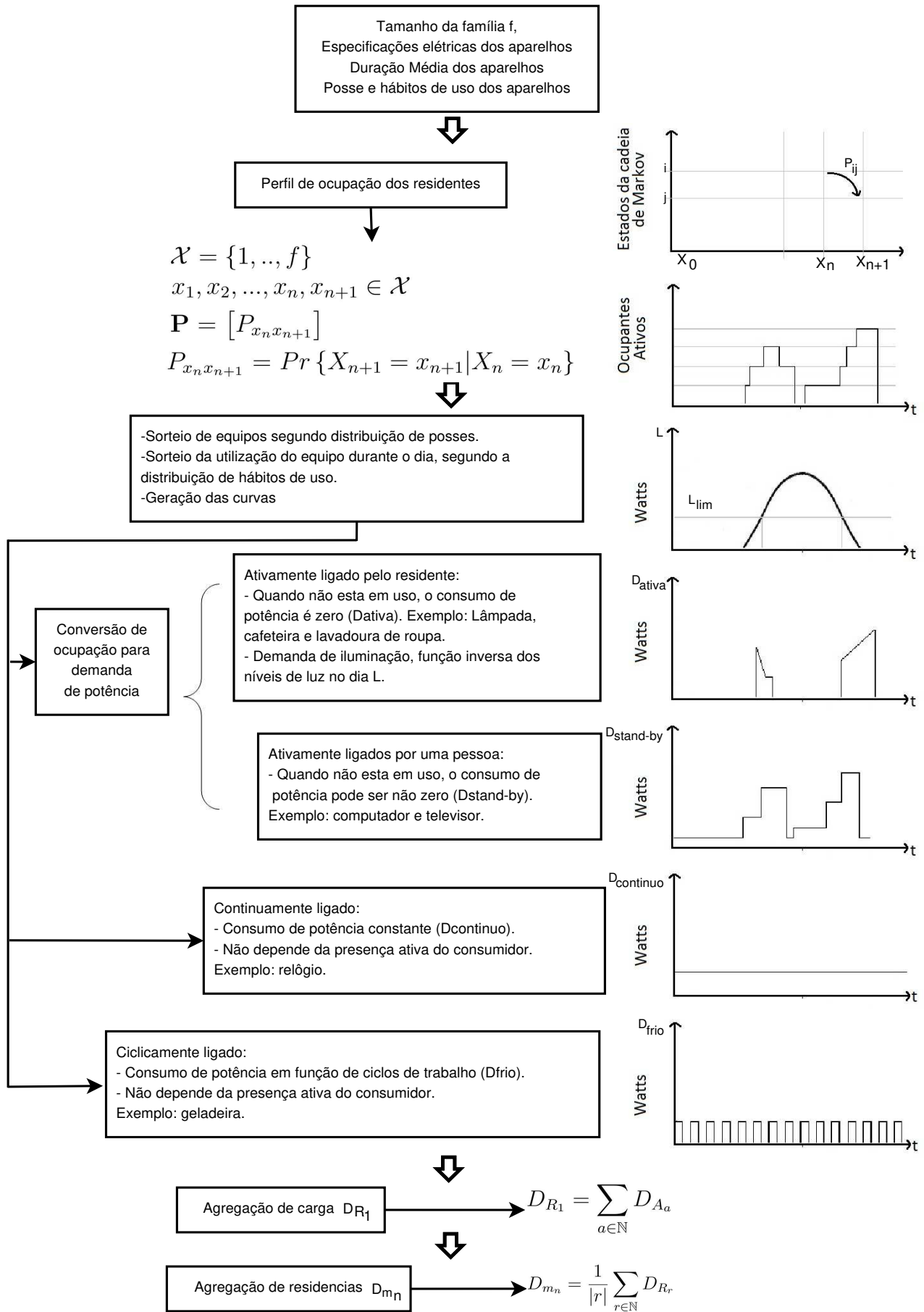


Figura 4.3: Esquema do processo de modelagem de demanda considerando a participação das atividades dos membros de uma residência.

Tabela 4.2: Especificações de aparelhos domésticos residenciais - Modelo 2.

Eletrodoméstico	Potência média (Watts)	Potência standby (Watts)	Duração média (Minutos)	Demora novo ciclo (Minutos)	Posse
Televisor	90	3	73	0	0,971
VCR/DVD	34	2	73	0	0,869
Receptor de Televisão	27	15	73	0	0,934
Micro-onda	1250	2	30	0	0,300
Liquidificador	200	0	1	0	0,820
Batedeira	100	0	5	0	0,470
Lava Roupa	1500	1	80	0	0,840
Chuveiro	3500	0	8	0	0,731
Ferro	1000	0	30	0	0,930
Aspirador de Pó	1000	0	20	0	0,100
Geladeira	200	0	18	36	0,960
Congelador	400	0	22	56	0,232
Relógio	0	2	0	0	0,900
Videocassete/DVD	15	2	60	0	0,800
Telefone	0	1	0	0	0,900
Roteador	100	9	60	0	0,900
Fax	37	3	31	0	0,200
Computador	250	5	300	0	0,708
Impressora	335	4	4	0	0,665
Cafeteira	1000	1	15	0	0,975
Ventilador	100	0	240	0	0,780
Vídeo Game	20	2	60	0	0,09
Ar Condicionado	4000	0	240	0	0,105
Aparelho de Som	20	2	180	0	0,74
Outros	2000	0	240	0	0,026

– Potência ativa e *stand-by*

Este grupo de aparelhos são os que dependem da presença ativa do consumidor para que estes sejam ligados. Após desligados, o consumo de potência pode ser zero ou não zero, este ultimo representado pela potência "stand-by". (i.e. televisor e lâmpada).

Um caso particular foi desenvolvido na demanda de iluminação, que é gerada a partir da conversão do perfil de ocupação e dos níveis de intensidade luminosa L , determinados a partir dos dados de irradiação solar mensais (Richardson & Thomson 2011, Widen, Nilsson & Wackelgard 2009).

Assim, enquanto exista ocupantes na casa, a demanda de luz é calculada da seguinte forma:

$$D_{ativa} = P_{min} \frac{L(n)}{L_{lim}} + P_{max} \left(1 - \frac{L(n)}{L_{lim}}\right), \quad L \leq L_{lim} \quad (4.3)$$

$$D_{ativa} = P_{min}, \quad L > L_{lim}$$

em que P_{min} e P_{max} são potências entre os diferentes tipos de lâmpadas comerciais, respectivamente, e L_{lim} é o nível de iluminação de referência.

– Potência contínua e por ciclos

São os aparelhos que estão constantemente ligados e não precisam da presença ativa do consumidor. Neste caso, o consumo de potência do aparelho pode ser constante no dia ou em função dos ciclos de trabalho, como por exemplo o relógio e geladeira, respectivamente. No caso de estar desligado é analisado o tempo de retardo para dar início a um novo ciclo, caso o aparelho trabalhe por níveis como a geladeira.

Com o propósito de exibir a aleatoriedade das curvas de cargas resultantes, a Fig 4.5 apresenta 10 perfis de cargas individuais para $T = 1 \text{ minuto}$.

4.2.3 Validação de Consumo de Energia Elétrica

A fim de validar os dados obtidos, é importante fazer uma comparação com as curvas reais medidas em campo.

É claro que as curvas de consumos individuais simuladas da Fig. 4.5 e Fig. 4.4, apresentam a aleatoriedade de consumo como é esperado e visualizado nas curvas reais da Fig. 4.6. Para estes, nota-se que existe maior variabilidade de carga nas horas de maior consumo, que seriam os intervalos de tempo onde os residentes permanecem ativos em casa e por consequência maior uso dos aparelhos. Isso é descrito pela curva de desvio padrão, que assume altos valores devido à diversidade de uso dos aparelhos, principalmente pelo chuveiro.

A fim de avaliar se os consumos de energias diárias dos residentes resultam em valores típicos que os residentes estão acostumados a pagar. Os pontos da Fig. 4.8, apresentam as energias consumidas diárias para um período de 60 *dias*, tanto para os dados simulados, como para dados reais.

Nota-se que no caso do primeiro modelo, em média a energia diária é de 7,62 kWh com um desvio de aproximadamente 1 kWh, apresentando pouca variação diária de consumo. Isso ocorre pelo fato de que a duração do uso de cada carga é considerada constante durante todos os dias, mesmo sendo aleatório o sorteio do uso do aparelho durante o dia.

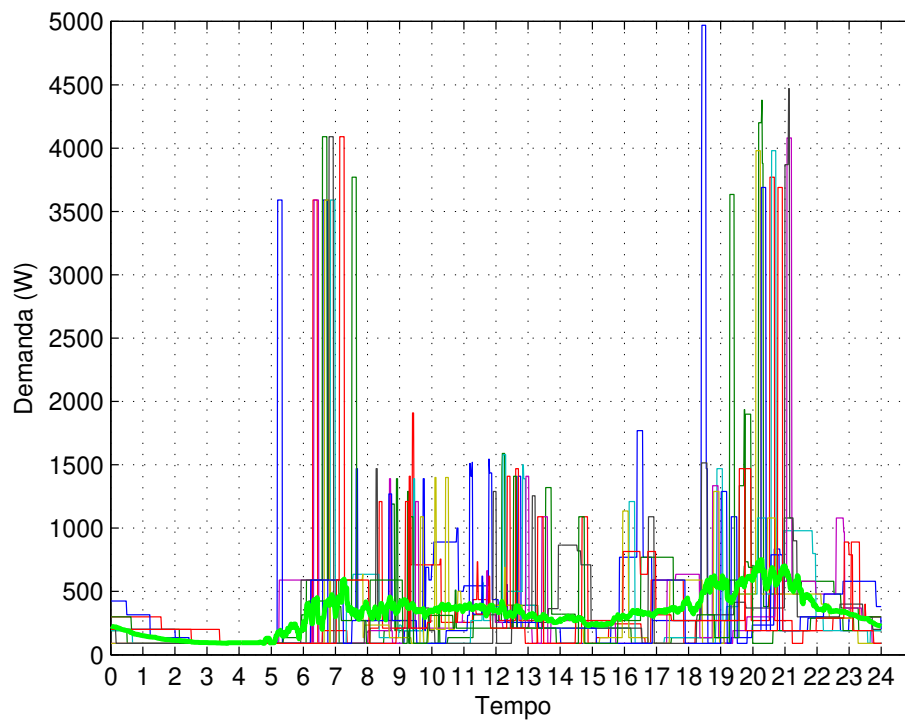


Figura 4.4: Dez exemplos de curvas diárias de consumo elétrico residenciais simuladas e a curva correspondente de desvio representando a diversidade de uso dos aparelhos. Resultados usando o primeiro modelo.

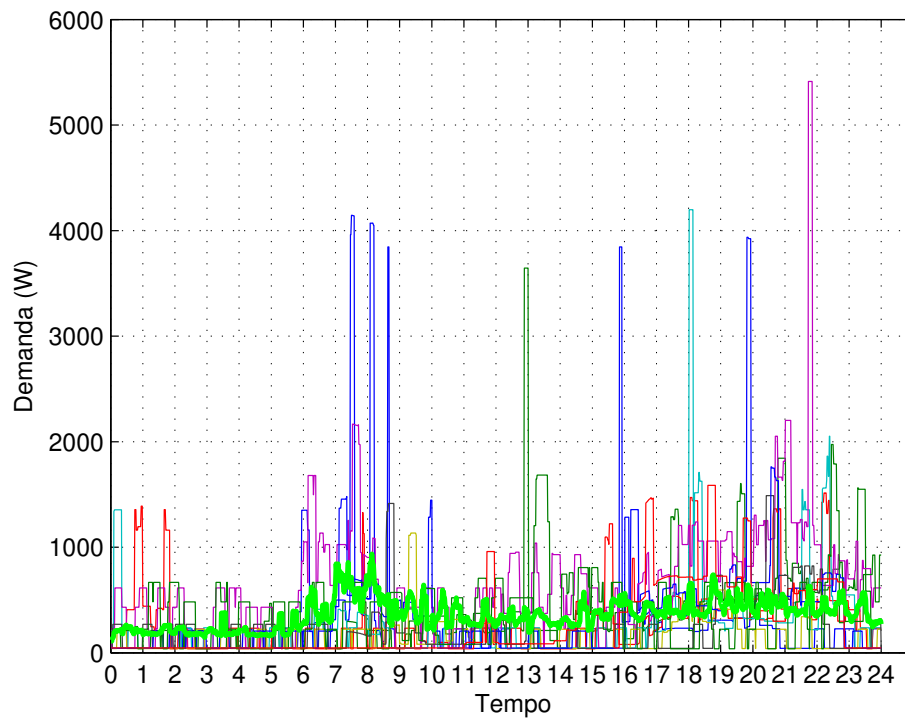


Figura 4.5: Dez exemplos de curvas diárias de consumo elétrico residencial simuladas e a correspondente curva de desvio representando a diversidade de uso dos aparelhos. Resultados usando o segundo modelo.

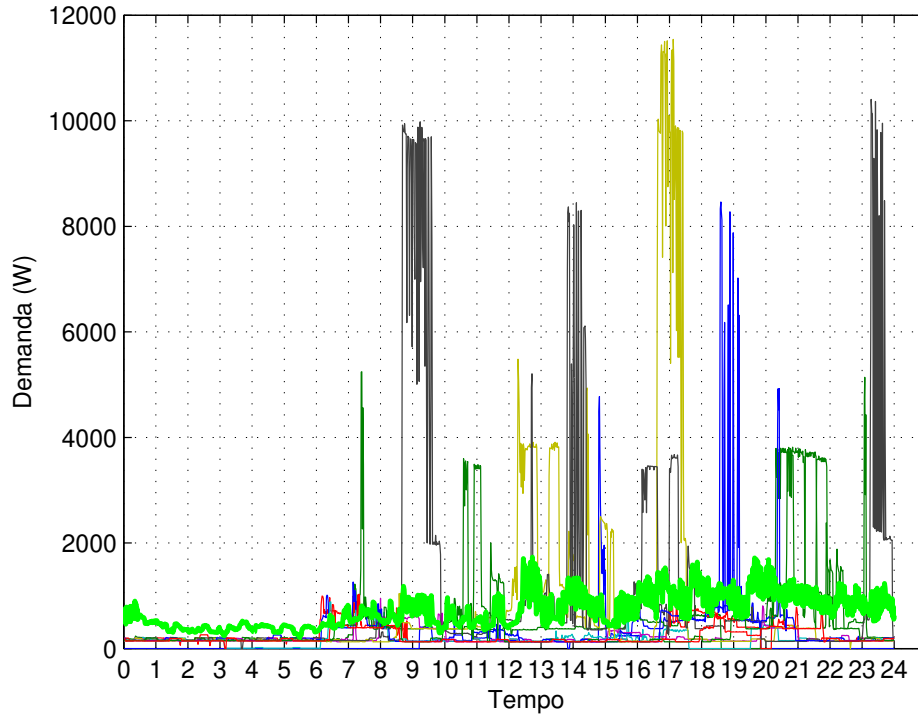


Figura 4.6: Dez exemplos de curvas diárias de consumo elétrico residencial **reais** e a correspondente curva de desvio representando a diversidade de uso dos aparelhos.

Por outro lado, no segundo modelo pode-se observar que em média a energia diária é de 8,8004 kWh com maior desvio de 3,0927 kWh/dia aproximadamente, cujo intervalo de energias resulta em valores típicos de um consumidor residencial real. Para o caso dos pontos reais, alguns dias apresentam consumos muito altos, e isso ocorre pelo fato de que os consumos, além de ser em aleatórios devido às cargas, são diferentes entre regiões, onde o uso de outros aparelhos de maior consumo, como o aquecedor é muito comum em países onde a maior parte do tempo há frio, à diferença do Brasil. Estes resultados são apresentados na tabela 4.3.

Além disso, as curvas de cargas representam um grupo residencial com faixa de consumo distribuído ao redor dos 366 W e com desvio de 162,4 W aproximadamente, como apresentado na Fig.4.9.

Tabela 4.3: Média e desvio padrão da energia consumida diária dados simulados e reais.

	Simulado - Modelo 1	Simulado - Modelo 2	Real
Média	7,62 kWh	8,80 kWh	9,01 kWh
Desvio Padrão	1 kWh	3,09 kWh	5 kWh

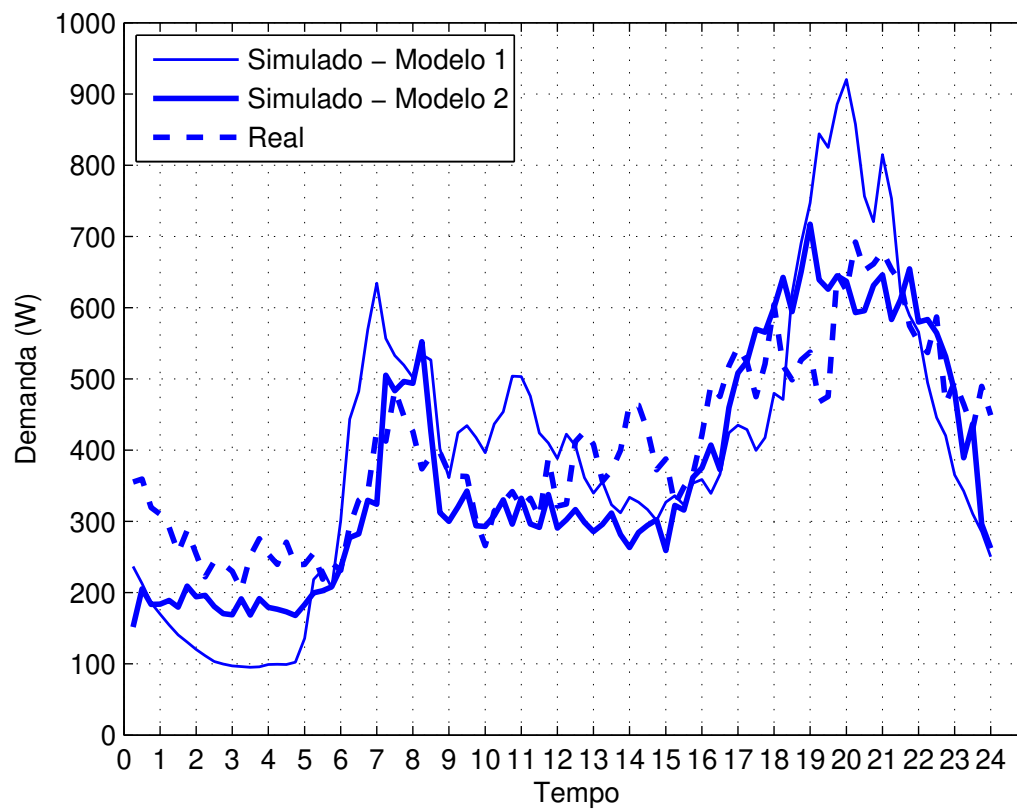


Figura 4.7: Curva de carga representativa do consumidor residencial.

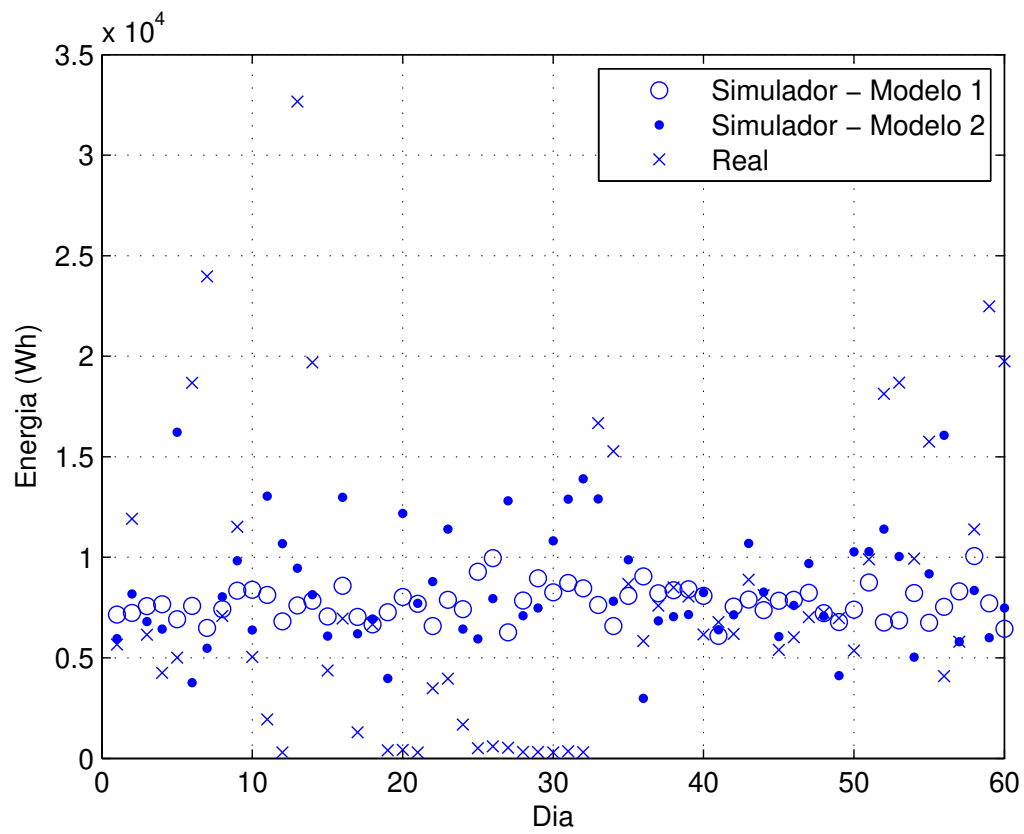


Figura 4.8: Energia consumida diária em 60 dias para dados simulados e reais.

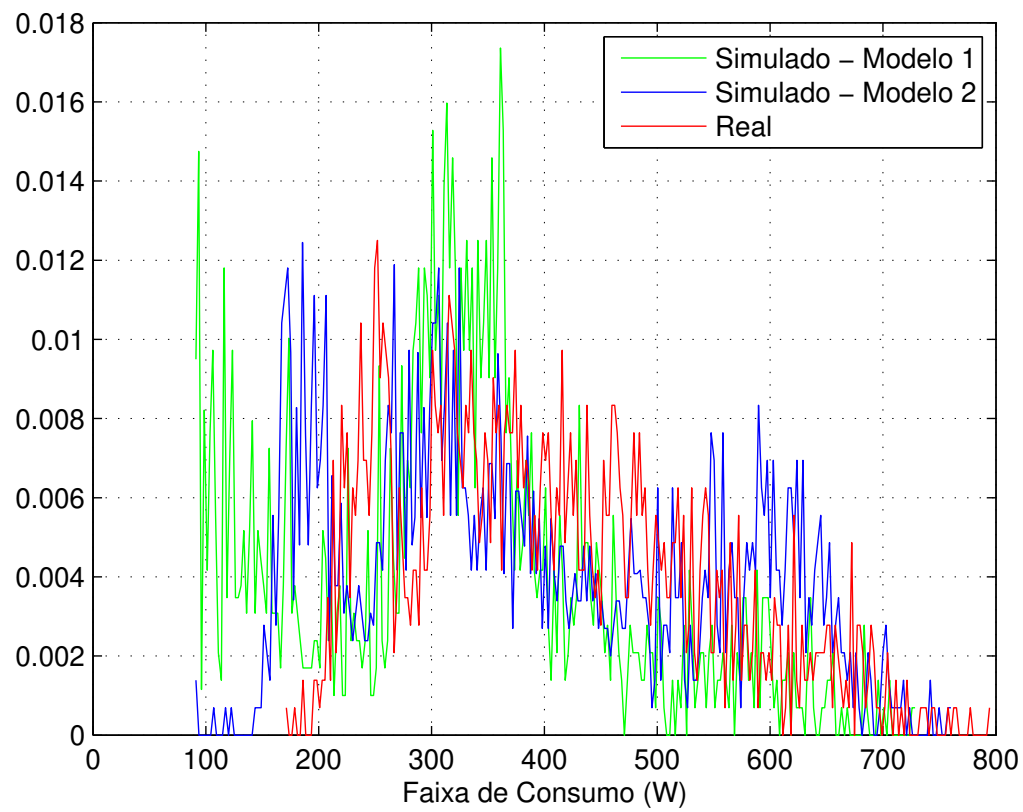


Figura 4.9: Distribuição do consumo de potências médias para dados simulados e reais.

4.3 Avaliação de Compressão

Tendo disponíveis os consumos para diferentes períodos de registro de dados e a correspondente distribuição estatística, é importante agora analisar os resultados obtidos do processamento de compressão dos dados usando o modelo proposto da Fig. 4.1, para diferentes cenários.

4.3.1 Distorção na Compressão da Informação de Consumo

Quando se avalia o desempenho dos esquemas de compressão com perdas som e imagens, está-se interessado na quantidade de informação que pode ser eliminado do sinal original, de tal forma que o sistema visual e auditivo humano não percebam a falta de informação (Tcheou et al. 2014). Quando a compressão é aplicada ao sinal de consumo elétrico, a distorção introduzida pela digitalização do sinal, leva a falhas e más interpretações, especificamente no sistema automatizados de engenharia para a análise de falhas.

Além disso, dado o fato de que a tarifação de consumo elétrico será diferenciada, ou seja, a cada valor de consumo corresponde uma tarifa dependendo do horário do dia, é importante que as informações recebidas na concessionária, resultantes da compressão, sejam as mais exatas possíveis para evitar perdas econômicas tanto para o consumidor como para a companhia.

As curvas da Fig. 4.10, descrevem o desempenho dos quantizadores PCM e DPCM aplicados aos consumos médios diários para períodos de registro variando de 1/900 Hz - 1 Hz. Em ambos dos casos, tanto para os simulados como para os reais, é evidente que quanto maior nível de quantização, menor é a distorção.

Se assumirmos 256 como o número de níveis de quantização, o período de $T = 15$ minutos representa menor distorção do que o período de 1 *segundo*, por exemplo. Isso é notado pelo fato de que o medidor a cada período, registra o consumo médio que foi acumulado nesse intervalo, portanto para períodos maiores, as curvas de consumo médio apresentam-se mais suaves do que as de menor período, o que leva a um erro menor na hora de fazer a quantização. Note na Fig. 4.10, que o esquema DPCM apresenta um ganho de SQNR em relação ao esquema PCM.

Embora as curvas tenham o mesmo comportamento na diminuição do SQNR para valores menores de T , pode-se notar que os resultados obtidos no segundo modelo de simulação são muito próximos aos dos dados reais. Estes resultados são apresentados na tabela 4.4 utilizando 256 níveis de quantização.

4.3.2 Taxa de Registro de Dados de Consumo

Normalmente as curvas diárias dos consumidores são medidas em períodos de 15 a 30 dias em média, com registros diários de potência média em intervalos de 1 hora, ou 15 minutos,

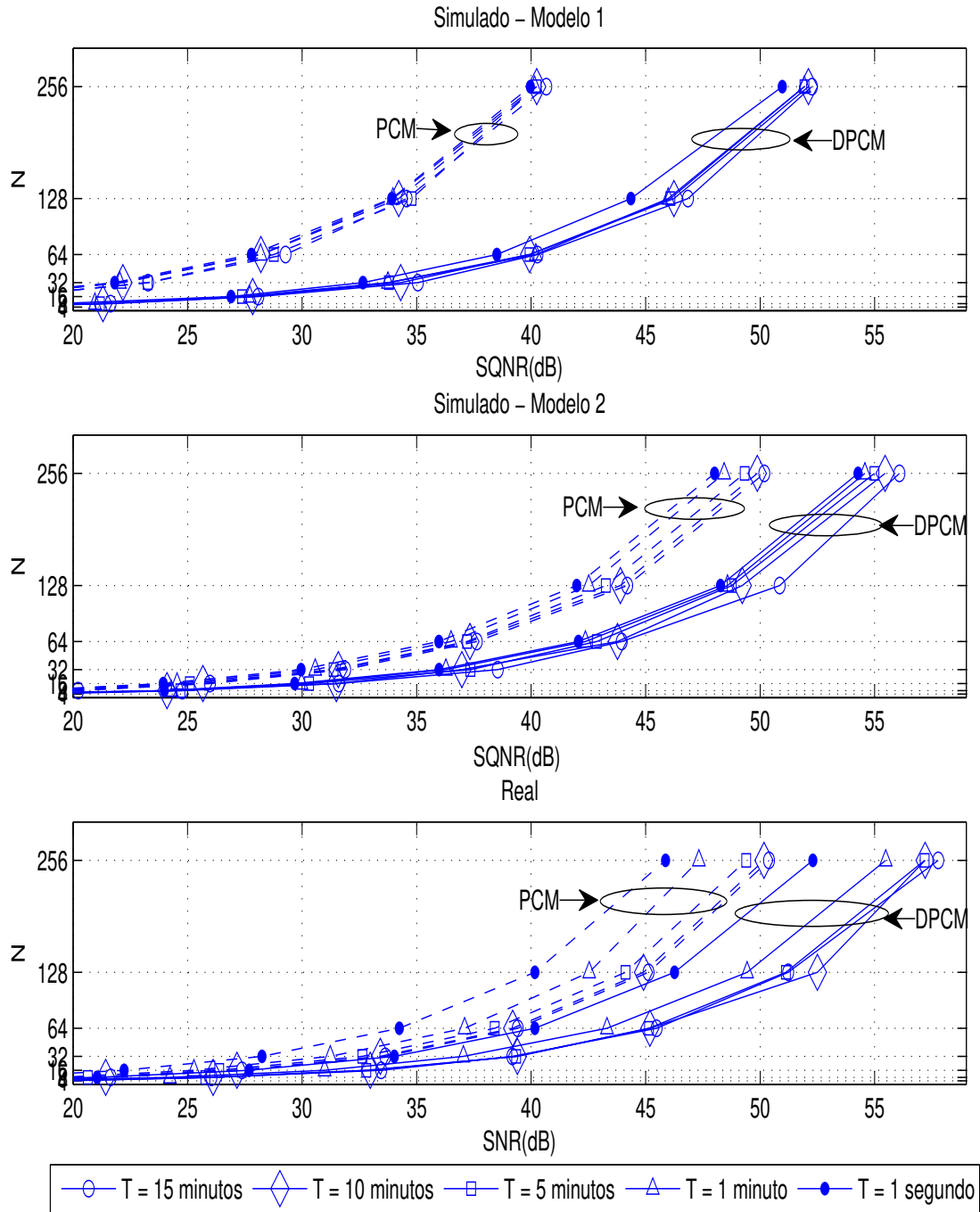


Figura 4.10: Relação sinal-ruído de quantização em PCM e DPCM, para dados de consumo simulados e reais, medidos em intervalos de $T = \{1\text{seg}, 1\text{min}, 5\text{min}, 10\text{min}, 15\text{min}\}$.

Tabela 4.4: Relação sinal-ruído de quantização para PCM e DPCM com 256 níveis de quantização, aplicados aos dados simulados e reais, com medições em períodos de $T = \{1 \text{ seg}, 1 \text{ min}, 5 \text{ min}, 10 \text{ min}, 15 \text{ min}\}$.

	SQNR PCM		
	Simulado - Modelo 1	Simulado - Modelo 2	Real
1 segundo	39.96 dB	48.01 dB	45.87 dB
1 minuto	40.15 dB	48.43 dB	47.32 dB
5 minutos	40.23 dB	49.32 dB	49.39 dB
10 minutos	40.25 dB	49.88 dB	50.17 dB
15 minutos	40.66 dB	50.18 dB	50.39 dB
	SQNR DPCM		
	Simulado - Modelo 1	Simulado - Modelo 2	Real
1 segundo	50.97 dB	54.28 dB	52.3 dB
1 minuto	51.92 dB	54.58 dB	55.48 dB
5 minutos	51.94 dB	54.99 dB	57.19 dB
10 minutos	52.11 dB	55.45 dB	57.21 dB
15 minutos	52.26 dB	56.08 dB	57.78 dB

representando uma taxa de amostragem de 1/900 Hz.

A análise em altas taxas de amostragem torna-se cada vez mais importante, tanto para as companhias elétricas como para os residentes, pois proporciona uma informação detalhada em tempo real sobre o consumo elétrico, permitindo detectar eventos de cargas que provocam determinados consumos, ajudando aos residentes a melhorar a gestão de seu consumo para um estilo de vida sustentável, mas também permite às companhias proporcionar programas de conservação de energia (Chen, Dai, Wang, Sahu, Naphade & Lu 2011).

Embora, os medidores podem amostrar em taxas ainda mais altas que 1 MHz, muitas implementações existentes tem escolhido o período de acumulação de 15 minutos ou ainda intervalos mais longos para assegurar a confiabilidade na transmissão. Por outro lado, é claro que há um custo em aumentar a taxa de amostragem, dado que quanto maior for esta taxa maior será a quantidade de dados a serem armazenados. Portanto, é evidente que existe um compromisso na escolha da taxa de amostragem.

A Fig.4.11, apresenta o comportamento do limite máximo de compressão medido em bits médios por amostra de medição (b/a), em função dos níveis de quantização (N) utilizados para digitalizar o sinal de tensão e corrente de consumo de energia elétrica média residencial para períodos de registro de dados variando de 1/900 Hz - 1 Hz.

Assumindo que o sinal é quantizado com 256 níveis e usando um período de registro de 15 minutos, consegue-se uma compressão de até 6,13 *bits/amostra* em média, quando o intervalo utilizado é de 1 segundo, o comprimento médio será de aproximadamente 7,40 *bits/amostra*.

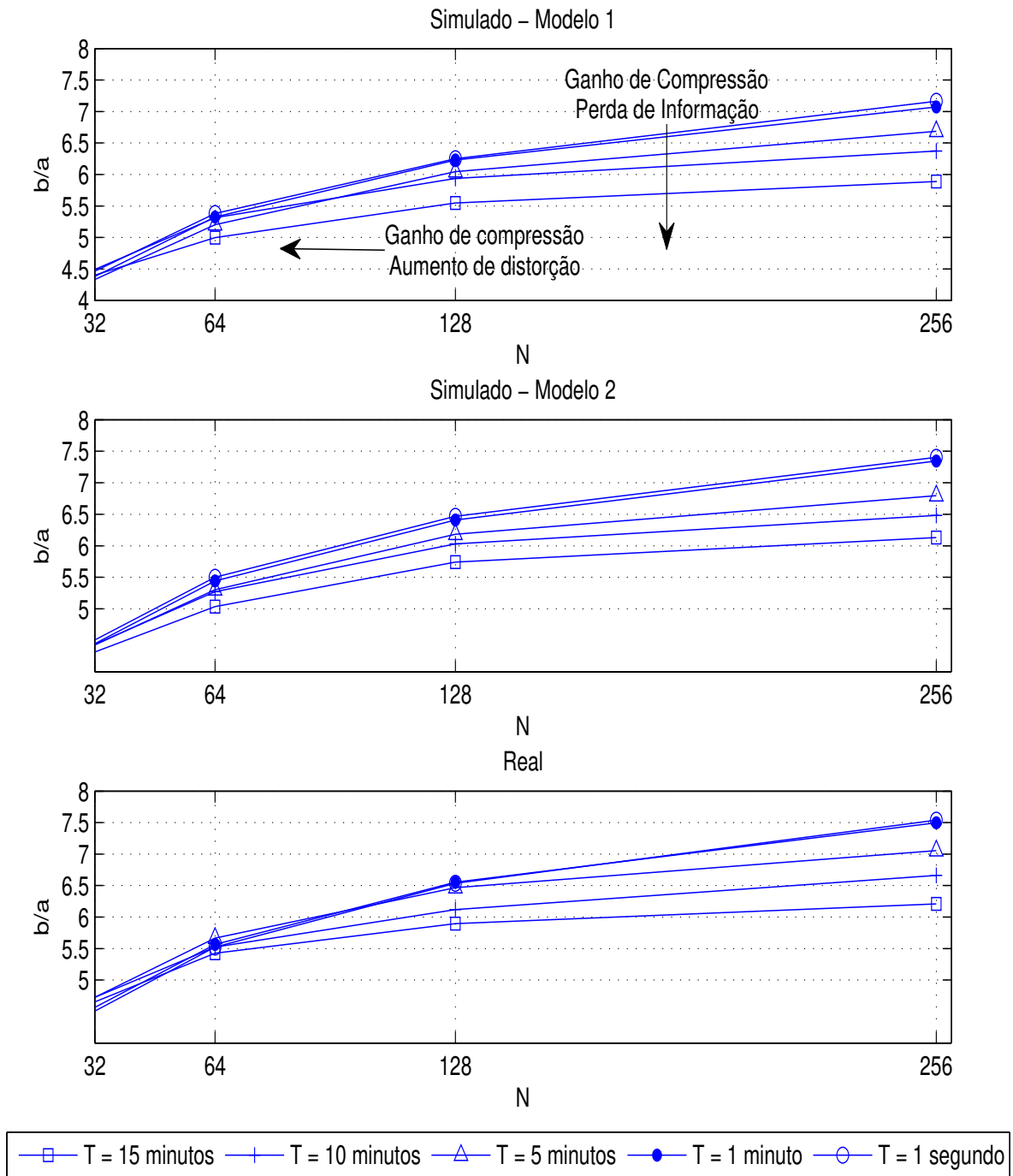


Figura 4.11: Comprimento médio em bits por amostra de medição de consumo diário para os dados simulados e reais, em função dos níveis de quantização e intervalos de registro $T = \{1\text{seg}, 1\text{min}, 5\text{min}, 10\text{min}, 15\text{min}\}$.

Percebe-se que uma maior taxa de compressão é alcançada em um período de amostragem de 15 *minutos*. No entanto, esses números não levam em conta a importância da informação sobre o consumo quando se utiliza uma alta taxa de amostragem, necessária para as distribuidoras de energia elétrica.

Se considerarmos a taxa de compressão e a qualidade da informação, pode-se assumir um período de registro ideal de 1 *minuto*, já que não há nenhuma alteração significativa na entropia quando a taxa é de $T = 1$ *segundo*. Em outras palavras, é desnecessário aumentar a taxa de amostragem acima desse número pois não há nenhum ganho na entropia. Conclui-se também que 15 *minutos* de período de amostragem implica em uma perda de detalhe na informação.

Estes resultados são apresentados na tabela 4.5 quando são usados 256 níveis de quantização. Assim, igual que na seção anterior, pode-se notar que os resultados obtidos do segundo modelo de simulação são muito próximos aos obtidos usando os dados reais.

Tabela 4.5: Comprimento médio em bits por medição de consumo diário para 256 níveis de quantização, aplicados aos dados simulados com medição em períodos $T = \{1\text{seg}, 1\text{ min}, 5\text{ min}, 10\text{ min}, 15\text{ min}\}$.

	Simulado - Modelo 1	Simulado - Modelo 2	Real
1 segundo	7.161 bits	7.405 bits	7.53 bits
1 minuto	7.073 bits	7.346 bits	7.497 bits
5 minutos	6.685 bits	6.794 bits	7.056 bits
10 minutos	6.37 bits	6.48 bits	6.66 bits
15 minutos	5.89 bits	6.132 bits	6.20 bits

4.3.3 PCM-Huffman e DPCM-Huffman

Dado que a codificação de Huffman é uma técnica de codificação ótima de comprimento variável, esta é aplicada neste modelo para avaliar seu desempenho junto com PCM e DPCM com relação ao limite teórico de compressão ou entropia de primeira ordem, usando a Eq. (3.15). Isso é apresentado nas curvas da Fig. 4.12 para períodos de $T = 15$ *minutos* e $T = 1$ *minuto* em dados simulados e reais, respectivamente.

Dos resultados, nota-se que um esquema não eficiente é conseguido quando a distribuição da fonte é uniforme, já que como N é uma potência de 2 ou $N = 2^b$, então b bits são necessários para a representação de cada nível. Assim, quando $N = 256$, no máximo são necessários 8 *bits* por amostra de consumo. Pode-se diminuir o número de bits por amostra quando se considera um processo com correlação, como será visto a seguir.

Conforme é observado, no melhor caso DPCM-Huffman reduz cerca de 10% quando o período de registro empregado é de $T = 15$ *min* com relação à compressão obtida usando PCM-Huffman.

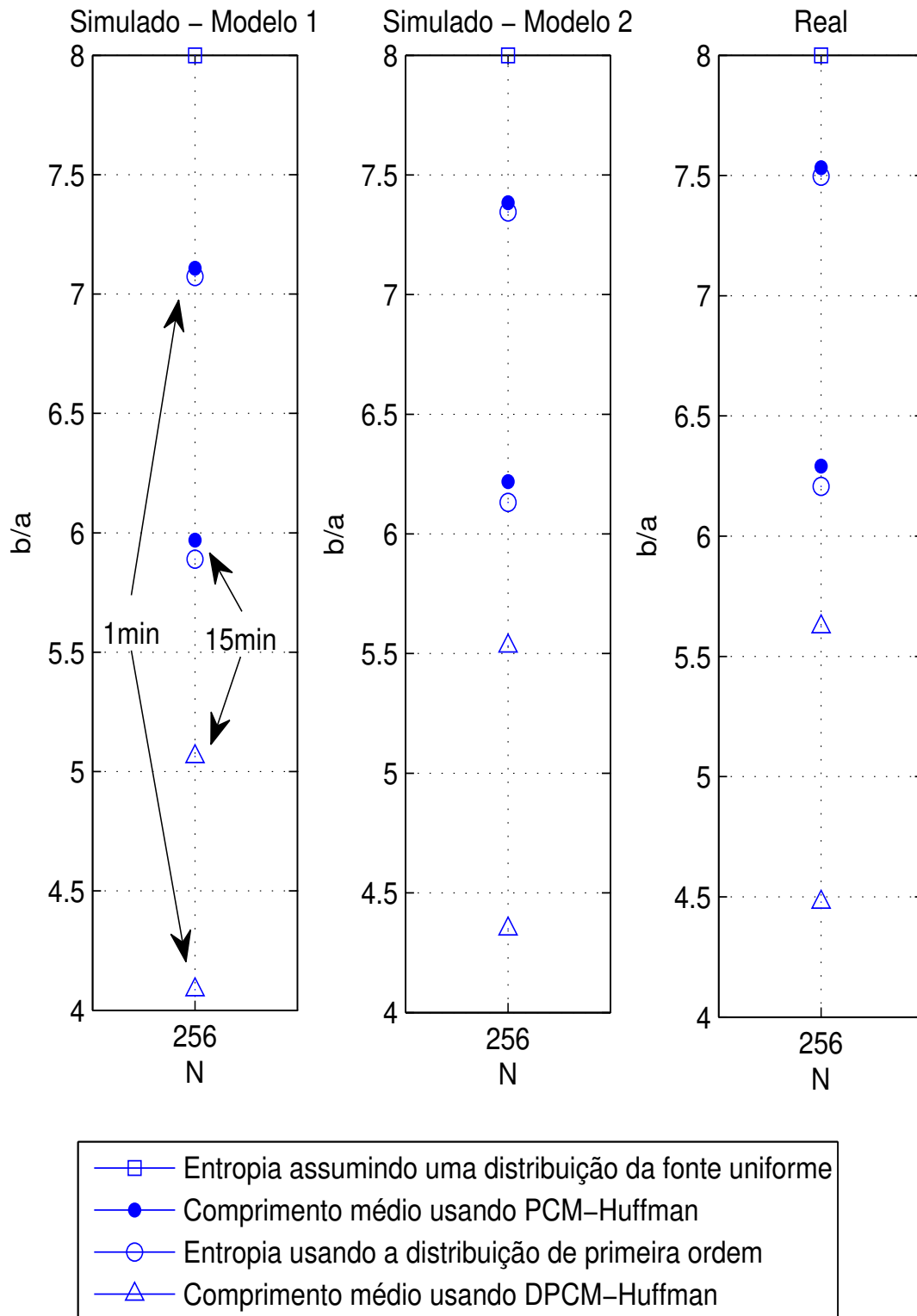


Figura 4.12: Comprimento médio em bits por medição de consumo diário para os dados simulados e reais com 256 níveis de quantização para uma fonte uniformemente distribuída e de primeira ordem, esquemas PCM-Huffman e DPCM-Huffman, em períodos de $T = 1$ minuto e 15 minutos.

Usando a mesma análise, para $T = 1 \text{ min}$, o limite de compressão é reduzido para 40%. Isso se deve ao fato de que DPCM quantiza as diferenças entre amostras sucessivas em vez das próprias amostras. Como as diferenças são de magnitude menor do que as amostras, requer-se um número maior de bits para quantizar a amplitude.

Nota-se também, como o código PCM-Huffman aproxima-se do limite de compressão teórico dado pela entropia de primeira ordem, como foi estabelecido na Eq. (3.13), descrito no capítulo anterior.

O comprimento médio por amostra é muito útil para estimar o número total de bits e, portanto, a memória de disco para armazenar a informação diária. Uma vez que temos $T = 15 \text{ minutos}$, em um dia serão necessários $24/0,15 = 96$ amostras por dia. Considerando 256 níveis de quantização e a utilização do esquema PCM-Huffman, o número total de bits será $6,21 \times 96 = 596,16$ bits em um dia. Isto parece um número insignificante, mas quando consideramos uma cidade com 17 milhões de pessoas como São Paulo, isso implicaria em quase 10,13 bilhões de bits em um dia, que é um número importante.

Se o mesmo raciocínio é realizado para um codificador DPCM-Huffman, obtemos 9,02 bilhões de bits. Se o período de amostragem for de $T = 1$ minuto, teríamos quase 180,7 e 106,5 bilhões de bits para PCM-Huffman e DPCM-Huffman, respectivamente. Estes números são apresentados na tabela 4.6, cujos dados ressaltam a importância desta análise.

Tabela 4.6: Numero médio de bits totais diários para representar o consumo residencial de uma região de 17 milhões de residentes assumindo 256 níveis de quantização para períodos de $T = \{1 \text{ min}, 15 \text{ min}\}$.

	T = 1 minutos		
	Simulado - Modelo 1	Simulado - Modelo 2	Real
PCM - Huffman	7.108 bits	7.383 bits	7.53 bits
Total bits/dia	174 Gb	180.7 Gb	184.33 Gb
DPCM - Huffman	4.086 bits	4.35 bits	4.478 bits
Total bits/dia	100.03 Gb	106.5 Gb	109.62 Gb
	T = 15 minutos		
	Simulado - Modelo 1	Simulado - Modelo 2	Real
PCM - Huffman	5.96 bits	6.21 bits	6.29 bits
Total bits/dia	9.72 Gb	10.13 Gb	10.26 Gb
DPCM - Huffman	5.03 bits	5.53 bits	5.62 bits
Total bits/dia	8.21 Gb	9.02 Gb	9.17 Gb

4.3.4 Ordem de Entropia

A fim de investigar o efeito da correlação entre as amostras, a taxa da entropia de segunda e terceira ordem foram calculadas de acordo com a Eq. (3.16) e Eq. (3.18), respectivamente.

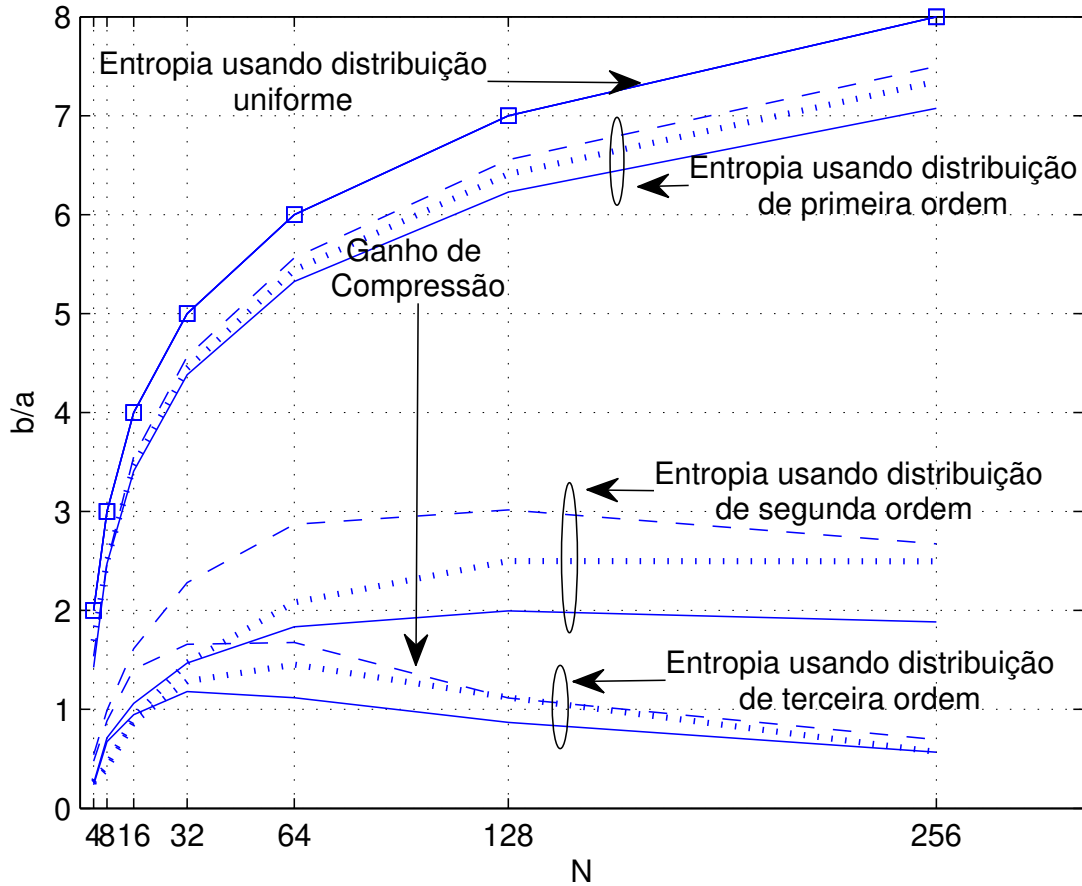


Figura 4.13: Comprimento médio em bits por amostra para consumo diário em função dos níveis de quantização para diferentes ordens de distribuição da fonte. Dados simulados e reais em $T = 1 \text{ minuto}$. Linha contínua representa "Simulado - Modelo 1", linha tracejada "Simulado - Modelo 2" e pontilhada "Real"

A Fig. 4.13 apresenta o desempenho destes limites de compressão ou entropias quando assume-se que a fonte pode ser representada estatisticamente por uma distribuição uniforme ou de ordem zero, de primeira ordem com amostras independentes, de segunda e terceira ordem com amostras condicionalmente dependentes das anteriores.

A tabela 4.7 resume a comparação entre os resultados dos limites de compressão e o desempenho dos esquemas de compressão para um período de registro de 1 minuto e 256 níveis de quantização.

Os resultados são muito interessantes, uma vez que mostram que maiores taxas de compressão podem ser alcançadas se a correlação entre amostras consecutivas é considerada. Como pode ser visto, para 256 níveis de quantização e período de registro de $T = 1 \text{ minuto}$, um código com somente 2,49 *bits/amostra* seria requerido, o que representa uma redução de 66.26% e 42.75%

quando comparados com os esquemas PCM-Huffman e DPCM-Huffman baseados em entropia de primeira ordem, respectivamente.

Tabela 4.7: Comprimento médio em bits por medição de consumo diário para 256 níveis de quantização e períodos de 1 minuto para os esquemas PCM-Huffman, DPCM-Huffman e processos de primeira e segunda ordem.

	Simulado Modelo 1	Simulado Modelo 2	Real
$\mathcal{H}$-Distribuição uniforme	8 bits	8 bits	8 bits
PCM - Huffman	7.108 bits	7.383 bits	7.53 bits
$\mathcal{H}_1$-Distribuição de primeira ordem	7.07 bits	7.346 bits	7.49 bits
DPCM - Huffman	4.086 bits	4.35 bits	7.478 bits
$\mathcal{H}_2$-Distribuição de segunda ordem	1.88 bits	2.49 bits	2.67 bits
$\mathcal{H}_3$-Distribuição de terceira ordem	0.566 bits	0.569 bits	0.69 bits

4.4 Resumo do Capítulo

Neste capítulo foram apresentados os resultados numéricos e comparações dos valores de compressão obtidos, considerando diversos cenários, a fim de validar as expressões analíticas do limite teórico de compressão.

Embora as duas modelagens empregadas são muito realistas, alcançando fatores característicos da curva de carga próximos aos valores reais medidos em campo, o desempenho do segundo modelo ficou mais próximo do que o primeiro modelo, quando comparados ao desempenho utilizando dados reais.

A avaliação da compressão levou em conta: a distorção provocada pela quantização da amplitude, a taxa de registro de dados, e a correlação entre as amostras. Os resultados mostraram que a compressão utilização DPCM é muito mais eficiente do que PCM, conforme o esperado.

Concluimos que os melhores esquemas de compressão exploram a correlação entre as amostras. Foi visto que usando uma modelagem de fonte de segunda ordem, consegue-se uma redução de 66,26% e 42,75% quando comparadas aos esquemas PCM-Huffman e DPCM-Huffman com amostras independentes, respectivamente.

Conclusões e Perspectivas

O trabalho consistiu em calcular o limite teórico de compressão dos dados de consumo residencial de energia média registradas em até um período de 1 *segundo*.

Elem disso, os esquemas PCM e DPCM combinado com a codificação de Huffman confirma ser muito eficiente na compressão sem perdas dos dados de consumo, com desempenho muito próximo ao limite teórico ou entropia de primeira ordem.

Foram simuladas as curva de demanda residencial baseados em dois modelos desenvolvidos neste trabalho. O segundo modelo mostrou-se mais realista, apesar de mais complexo, pois leva em conta as atividades dos membros das residências, posse de eletrodomésticos e hábitos de consumo dos aparelhos elétricos. Todos os resultados simulados apresentam uma boa concordância com os resultados usando os dados reais, como mostradas nas figuras relacionadas no capítulo anterior. No entanto, o segundo modelo consegue uma melhor aproximação.

Os resultados deste trabalho, representam um avanço relevante no âmbito da compressão de dados de consumos de demanda, uma vez que esta área está sendo recentemente estudada para o desenvolvimento da camada de comunicação em *Smart Metering*. Sendo assim, acreditamos que os resultados deste projeto podem ser uma ferramenta útil para prever os efeitos de gerenciamento de carga de dados.

5.1 Soluções futuras

Visto que na infraestrutura de comunicações, uma variedade de monitores e medidores inteligentes serão implantados e distribuídos no sistema de potência para permitir um monitoramento em tempo real, é fundamental verificar a massa de dados produzidos pelos equipamentos e seu impacto na infraestrutura de comunicações. Assim, o trabalho pode-se estender analisando que tipo de tecnologia será empregada para suportar essa quantidade de dados, em termos de largura de banda, segurança e padrões regulatórios.

Adicionalmente, pode-se ampliar o trabalho, com a aplicação de outras técnicas de compressão de dados baseadas em transformadas.

Trabalho Publicado

Durante o desenvolvimento do trabalho de tese, desde fevereiro de 2012 a agosto de 2014, foi publicado o seguinte trabalho:

- A. Carolina Flores, G. Fraidenraich, "Data Compression and Sampling Period of Daily Energy Consumption", Simpósio da SBrt - Sociedade Brasileira de Telecomunicações, Fortaleza, Ceará, Set. 2013.

Bibliografia

- Ardakanian, O., Keshav, S. & Rosenberg, C. (2011). Markovian Models for Home Electricity Consumption, *GreenNet* pp. 31–36.
- Capasso, A., Grattieri, W., Lamedica, R. & Prudenzi, A. (1994). A Bottom-Up Approach to Residential Load Modeling, *IEEE Transactions on Power Systems* **9**(2): 957–964.
- Chen, F., Dai, J., Wang, B., Sahu, S., Naphade, M. & Lu, C. (2011). Activity Analysis Based on Low Sample Rate Smart Meters, *KDD'11*.
- Covers, T. & Thomas, J. (2006). *Elements of Information Theory*, Vol. 2nd ed.
- EFLUL (2013). Tabela de Consumo dos Aparelhos, *Disponível em:* <http://www.eflul.com.br/consumidores/tabela-de-consumo>.
- Energimyndigheten (2013). Measuring Household Electricity at Individual Equipment Level, *Disponível em:* <http://www.energimyndigheten.se/en/Facts-and-figures1/Improved-energy-statistics-in-buildings/Measuring-household-electricity-at-individual-equipment-level/>.
- EXAME (2013). O Brasil na Onda das Smart Grids, *Disponível em:* <http://exame.abril.com.br/revista-exame/edicoes/1040/noticias/o-brasil-na-onda-das-smart-grids>.
- Gallager, R. (1976). *Information Theory and Reliable Communication*.
- Gellings, C. (2012). *The Smart Grid: Enabling Energy Efficiency and Demand Response*, Vol. 1.
- Gerek, O. & Ece, D. (2008). Compression of Power Quality Event Data Using 2D Representation, *Electric Power Systems Research* **78**: 1047–1052.
- Harney, A. (2009). Smart Metering Technology Promotes Energy Efficiency for a Greener World, *Analog Dialogue* **43-01**(01): 1–3.
- IEEE SmartGrid (2013). Smart Grid Conceptual Model, *Disponível em:* <http://smartgrid.ieee.org/ieee-smart-grid/smart-grid-conceptual-model>.
- Jardini, J. (1999). Curva de Carga, **2**(1): 12–15.

- Kay, S. (2012). *Intuitive Probability and Random Processes Using Matlab*.
- Kolter, J. & Johnson, M. (2011). REDD: A Public Data Set for Energy Disaggregation Research, *SustKDD*.
- Kraus, J., Robiska, T. & Bubla, V. (2009). Looseless Encoding and Compression Algorithms Applied on Power Quality Datasets, *IEEE CIRED* **1**: 1–4.
- Li, H., Mao, R., Lai, L. & Qiu, R. (2010). Compressed Meter Reading for Delay Sensitive and Secure Load Report in Smart Grid, *IEEE* pp. 114–119.
- Lorio, F. & Magnago, F. (2004). Analysis of Data Compression Methods for Power Quality Events, *IEEE Power Eng.* **1**: 504–509.
- Mehta, K. & Russell, B. (1989). Data Compression for Digital Data from Power Systems Disturbances: Requirements and Technique Evaluation, *IEEE Transactions on Power Delivery* **4**(3): 1683–1688.
- Momoh, J. (2012). *Smart Grid Fundamentals of Design and Analysis*, Vol. 1.
- Nunes, A., Da Costa, P., De Oliveira, R., Goretti, M., Papa, J. & Sikiti, D. (2010). Algoritmos para Estimar Curvas de Cargas a partir de Padrões de Hábitos de Consumo, *9th Brazilian Conference on Dynamics Control and their Applications* pp. 775–781.
- Proakis, J. & Salehi, M. (1997). *Contemporary Communication Systems*.
- PROCEL & ELETROBRÁS (2005). *Avaliação do Mercado de Eficiência Energética do Brasil*.
- PROCELINFO (2013). Sistema de Informações de Posses de Eletrodomésticos e Hábitos de Consumo, *Disponível em: <http://www.procelinfo.com.br/Sinpha/CurvaCarga.asp>*.
- Richardson, I. & Thomson, M. (2011). Integrated Simulation of Photovoltaic Micro-Generation and Domestic Electricity Demand: a one-minute resolution open-source model, *Loughborough University Institutional Repository*.
- Richardson, I., Thomson, M. & Infield, D. (2008). A High Resolution Domestic Building Occupancy Model for Energy Demand Simulations, *Energy and Buildings* **40**: 1560–1566.
- Rigwelki, M., Renner, C., Reinhardt, A., Weigel, A. & Turau, V. (2012). The Hitchhiker’s Guide to Choosing the Compression Algorithm for your Smart Meter Data, *IEEE ENERGYCON* pp. 998–1003.
- Romano, R. & Sípoli, P. (2013). Infraestruturas Avançadas de Medição, *CPqD* **9**(1): 37–46.
- Tcheou, M., Lovisol, L., Ribeiro, M., Da Silva, E., Rodrigues, A., Romano, J. & Diniz, P. (2014). The Compression of Electric Signal Waveforms for Smart Grids: State of the Art and Future Trends, *IEEE Transactions on Smart Grid* **5**(1): 291–302.

- Tsuji, K., Sano, F., Ueno, T. & Saeki, O. (n.d.). Bottom-Up Simulation Model for Estimating End-Use Energy Demand Profiles in Residential Houses, pp. 342–355.
- Widen, J., Nilsson, A. & Wackelgard, E. (2009). A Combined Markov-Chain and Bottom-Up Approach to Modelling of Domestic Lighting Demand, *Energy and Buildings* **41**: 1001–1012.
- Widen, J. & Wackelgard, E. (2010). A High Resolution Stochastic Model of Domestic Activity Patterns and Electricity Demand, *Applied Energy* **87**: 1880–1892.
- Wijaya, T., Eberle, J. & Aberer, K. (2013). Symbolic Representation of Smart Meter Data, *EDBT/ICDT'13* pp. 1–7.
- Zhang, D., Bi, Y. & Zhao, J. (2009). A New Data Compression Algorithm for Power Quality Online Monitoring, *SUPERGEN'09* pp. 1–4.