

Identificando região de gravação com base na ENF

Alisson S. Barreto, Felipe A. M. Alcântara, Rafael J. Lima, Carlos A. D. Pinto, João M. Gris, José A. Apolinário Jr. e Felipe A. Caetano de Bastos

Resumo—Uma gravação de áudio pode estar corrompida por um sinal da rede elétrica, a ENF (do inglês *Electrical Network Frequency*). A assinatura do sinal de rede é dependente da região—grande área ou país—de maneira que pode ser explorada para descobrir a localização de gravações de mídia. Neste trabalho, construímos métodos de prever estas localizações utilizando ferramentas consolidadas na área de análise de dados, os classificadores SVM e Random Forest. Após apresentarmos o problema, comparamos os resultados obtidos na aplicação destas técnicas a um mesmo conjunto de dados. É discutida também a possibilidade da amostra não pertencer a nenhuma das regiões.

Palavras-Chave—Frequência da rede elétrica, análise forense, rede elétrica, aprendizagem de máquina.

I. INTRODUÇÃO

Este trabalho de iniciação científica tem por objetivo implementar um sistema de classificação da região de gravação de um sinal de áudio. Ele também objetiva avaliar os parâmetros deste sistema, tais como os utilizados em [1], extraídos da frequência da rede elétrica (ENF). As redes de distribuição de energia elétrica possuem mecanismos de controle que limitam a variação da frequência da rede, cujo valor nominal é de 50Hz ou 60Hz. Buscou-se verificar a influência de um novo parâmetro, a entropia, e do caso da amostra não pertencer a nenhuma das regiões anteriores (NDA).

II. AQUISIÇÃO DOS SINAIS

Os sinais utilizados neste trabalho foram obtidos pela participação na Signal Processing Cup 2016, cujos sinais vieram separados em nove classes de distintas ENFs, e pela gravação de sinais análogos aos da competição pela equipe. Todos os sinais estavam no formato .wav e com frequência de amostragem igual a 1kHz.

III. EXTRAÇÃO DE CARACTERÍSTICAS DOS SINAIS

Para que as características ou *features* pudessem ser obtidas, extraímos primeiramente a ENF de cada sinal. Para tal, uma análise espectral indicou facilmente a frequência nominal da rede (50Hz ou 60Hz) e se o sinal era proveniente de uma gravação direta da rede elétrica (POWER) ou se tratava-se de um sinal de áudio corrompido com a ENF (AUDIO). Em ambos os casos, após tal decisão prosseguíamos estimando a ENF por meio do algoritmo DFT1 [2], [3]. Após a extração da ENF de todos os sinais, foram extraídas características (*features*) mostradas na Tabela I, as de [1] com a adição da entropia.

A. S. Barreto, R. J. Lima, J. A. Apolinário Jr. e F. A. C. Bastos, SE/3. F. A. M. Alcântara, C. A. D. Pinto J. M. Gris, SE/8, Instituto Militar de Engenharia (IME), Rio de Janeiro-RJ, Brasil, E-mails: alisson.de.s.b@gmail.com, carlos.adpinto@gmail.com, felipeaugusto1291@gmail.com, joamarcosgris@hotmail.com, rafael.ljosino1@hotmail.com, apolin@ieee.org, fcaetano@ime.eb.br.

A. Sinais de rede elétrica

Ao se gravar diretamente de uma tomada elétrica, pode-se observar, pelo espectro de frequências do sinal, tal como o exemplo da Figura 1 (b), a ENF e harmônicos ímpares dela com elevada relação sinal-ruído, SNR. De tal SNR, usualmente observada em gravações desse tipo, conclui-se que as ENF's desse tipo podem ser caracterizadas por apenas um dos harmônicos presentes, aquele de maior intensidade.

B. Sinais de áudio

Os sinais de áudio, por apresentarem em suas composições baixas frequências, tornaram difícil a identificação dos possíveis harmônicos da ENF e, ainda que identificados, por apresentarem ruído considerável na faixa de frequências da ENF, eram de baixa confiabilidade. Assim sendo, fez-se necessário que, para os sinais de áudio, as ENF's fossem estimadas por vários harmônicos ponderados por suas SNR's a fim de que possíveis erros devidos ao ruído do sinal fossem minimizados [4]. A Figura 1 (a) ilustra o espectro de um sinal de áudio corrompido pela ENF enquanto que a Figura 1 (c) mostra a estimação da ENF por meio da segunda harmônica e a necessidade de uma filtragem (filtros mediana e passa-baixas) na ruidosa sequência obtida para que ela se aproxime, após uma divisão por dois, da sequência obtida com o sinal POWER. Note na Figura 1 (d) que ambas ficam parecidas quando efetuamos a filtragem.

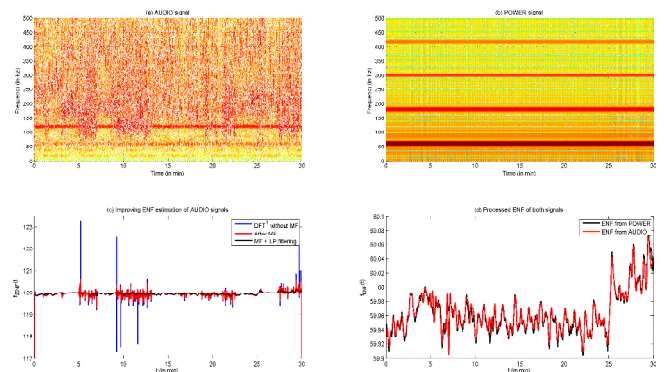


Fig. 1. ENF de sinais tipos POWER e AUDIO gravados simultaneamente.

IV. CLASSIFICADORES

Foram utilizados dois classificadores, o Support Vector Machines e o Random Forest. Ambos baseados no mesmo critério de classificação para o caso de NDA. A abordagem utilizada para este caso fez uso do vetor médio para cada classe e da maior distância existente entre uma amostra pertence a classe

TABELA I
Features USADAS NESTE TRABALHO

Índice do vetor	Feature(s)
1	Valor Médio da ENF
2	log(variância) da ENF
3	log(faixa de valores) da ENF
4	log(variância) da aproximação após análise wavelet de nível L=9
5-13	log(variância) de sinais de L=9 níveis de detalhes computados pela análise wavelet da mais grosseira à mais fina
14-15	Parâmetros a_1 e a_2 do modelo AR(2)
16	log(variância) do sinal de inovação após modelagem AR(2)
17	Entropia

TABELA II
DESEMPENHO DO SVM

Sem Entropia			
Treinamento	Teste	Precisão	Precisão com NDA
Power	Power	61.58 %	60.59 %
Power	Audio	43.33 %	43.33 %
Audio	Power	32.03 %	24.94 %
Audio	Audio	40.00 %	33.33 %
Com Entropia			
Treinamento	Teste	Precisão	Precisão com NDA
Power	Power	58.62 %	60.59 %
Power	Audio	43.33 %	43.33 %
Audio	Power	29.34 %	19.07 %
Audio	Audio	43.33 %	46.67 %

TABELA III
DESEMPENHO DO RANDOM FOREST

Sem Entropia			
Treinamento	Teste	Precisão	Precisão com NDA
Power	Power	84.75 %	83.53 %
Power	Audio	48.88 %	53.33 %
Audio	Power	42.35 %	43.53 %
Audio	Audio	61.11 %	50.00 %
Com Entropia			
Treinamento	Teste	Precisão	Precisão com NDA
Power	Power	84.14 %	83.53 %
Power	Audio	48.88 %	55.55 %
Audio	Power	45.00 %	42.94 %
Audio	Audio	55.55 %	44.44 %

e o vetor médio da mesma. Ao obter-se uma classificação sem o NDA, verifica-se se a distância entre a amostra classificada e a classe correspondente é maior que $1,5 \times$ máxima distância. Em caso positivo, a amostra é classificada como NDA.

A. Support Vectors Machine (SVM)

Utilizou-se como ferramenta na implementação do SVM o MATLAB[®] em conjunto com a biblioteca LIBSVM [1]. O SVM implementado neste trabalho utiliza em seu processo de classificação uma matriz de decisão, onde cada elemento dessa matriz é um classificador binário entre duas das classes possíveis. As classes menos frequentes na matriz ao tomar todos os classificadores serão eliminadas. Ao chegar em um estado onde todas as classes possuem a mesma frequência, é escolhida a classe com maior acurácia dentre as classes empatadas.

B. Random Forests

A implementação desse classificador foi baseada no pacote Python Scikit-learn. Esse pacote implementa a versão original do algoritmo descrito por [5]. Ele funciona construindo um grande número de árvores não correlacionadas. Após isso, as árvores votam pela classe mais popular [6].

V. RESULTADOS

As Tabelas exprimem os resultados encontrados para a implementação dos dois classificadores para a base de dados. Para cada tipo de classificador foram executadas quatro versões: com ou sem NDA e com ou sem Entropia. O Random Forest teve desempenho superior ao SVM em praticamente todos os casos de teste. Uma hipótese para justificar isso é que o Random Forest não precisa de grandes ajustes de parâmetros para funcionar bem, enquanto o SVM possui uma série de decisões que precisam ser tomadas, incluindo a escolha do Kernel e valores de parâmetros. Os valores calculados de entropia resultaram em números bastante próximos entre si, não contribuindo para a discriminação entre classes ou para caracterizar uma delas. A presença da escolha como NDA tendeu a piorar a classificação pois as distâncias entre features de power e de audio são grandes o suficiente para ocorrer um erro dessa classificação.

VI. CONCLUSÕES

A presença da Entropia se fez pouco notável e um tanto inconclusiva. A classificação com NDA mostrou uma tendência a piorar a taxa de acerto da classificação. A ferramenta Random Forest se mostrou mais adequada à presente aplicação.

AGRADECIMENTOS

Este trabalho foi parcialmente financiado pelo CNPq (Projeto Universal, Processo nr. 456901/2014-0) e pela CAPES (Projeto Prodefesa, Processo nr. 23038.009094/2013-83).

REFERÊNCIAS

- [1] A. Hajj-Ahmad, R. Garg, and M. Wu, "ENF-based region-of-recording identification for media signals," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 6, pp. 1125–1136, June 2015.
- [2] M. Desainte-Catherine and S. Marchand, "High-Precision Fourier Analysis of Sounds Using Signal Derivatives," *Journal of Audio Engineering Society*, vol. 48, no. 7/8, pp. 654–667, July/August 2000.
- [3] D. P. N. Rodríguez, J. A. Apolinário Jr., and L. W. P. Biscainho, "Audio authenticity: Detecting ENF discontinuity with high precision phase analysis," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 534–543, Sept. 2010.
- [4] A. Hajj-Ahmad, R. Garg, and M. Wu, "Spectrum combining for enf signal estimation," *IEEE Signal Processing Letter*, vol. 20, no. 9, pp. 885–888, September 2013.
- [5] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001. [Online]. Available: <http://dx.doi.org/10.1023/A:1010933404324>
- [6] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. Springer, 2008.