

Self-Organizing Fuzzy Rule-Based Approach for Dealing with the Classification of Indoor Environments for IoT Applications

Ualison R. F. Dias¹, Michel B. Hell²,
Álvaro A. M. de Medeiros¹, Daniel D. Silveira¹, Eduardo P. de Aguiar¹

¹Programa de Pós-Graduação em Engenharia Elétrica,
Faculdade de Engenharia – Universidade Federal de Juiz de Fora (UFJF)
Caixa Postal 36036-900 – Juiz de Fora – MG – Brazil

²Departamento de Circuitos Elétricos,
Faculdade de Engenharia – Universidade Federal de Juiz de Fora (UFJF)
Caixa Postal 36036-900 – Juiz de Fora – MG – Brazil

{ualison.dias,eduardo.aguiar}@engenharia.ufjf.br

Abstract. *Nowadays, a great part of the sensors used in the Internet of Things uses wireless technology in order to facilitate the construction of sensor networks. In this sense, the classification of the type of environment in which these sensors are located plays an important role in the performance of these sensor networks, since it leads to efficient power consumption when operating the deployed IoT sensors. Thus, this work presents an approach based on Self-Organizing Fuzzy Classifiers applied to indoor multipath environment classification from real-time measurements of the radiofrequency signal of a real wireless sensor network. Experimental results show that the proposed approach get high performance with low computational effort in the solution of the proposed problem.*

Resumo. *Atualmente, grande parte dos sensores utilizados em Internet das Coisas adota tecnologia sem fio, a fim de facilitar a construção de redes de sensoriamento. Neste sentido, a classificação do tipo de ambiente no qual estes sensores estão localizados exerce um importante papel no desempenho de tais redes de sensoriamento, uma vez que pode ser utilizada na determinação de níveis mais eficientes de consumo de energia dos sensores que as compõe. Assim, neste trabalho é apresentada uma abordagem baseada em Classificadores Fuzzy Auto-organizáveis para a classificação de ambientes internos a partir de medições em tempo real do sinal de radiofrequência de uma rede de sensoriamento sem fio em um ambiente real. Os resultados experimentais apresentados mostram que a abordagem proposta obteve alto desempenho com baixo custo computacional na solução do problema apresentado.*

1. Introdução

Com o advento da Internet das Coisas (IoT), as redes de sensores sem fio vêm desempenhando um papel cada vez mais importante em várias áreas de aplicação, tais como: saúde, agricultura, monitoramento de ambiente, medição inteligente, dentre tantas outras. Diversos tipos de sistemas de comunicação sem fio vêm sendo utilizados na construção de redes de sensores sem fio para aplicações de IoT, como nos equipamentos de comunicações via infravermelho, bluetooth, comunicação via micro-ondas, comunicação via satélite, e radiofrequência (RF). Neste contexto, redes de sensores baseadas em comunicação via RF vêm sendo largamente utilizadas. Isso se deve, principalmente, pelo fato de possuir grande eficiência energética, possibilidade de miniaturização de componentes (como possibilitado por sensores MEMS) e tecnologias computacionais incorporadas [Tiwari et al. 2015].

Com a maior mobilidade trazida pela ausência de conexões físicas em redes de sensores sem fio, esses dispositivos podem ser utilizados em diversos ambientes. Em muitos casos, a determinação do tipo de ambiente, onde o sensor está localizado desempenha um papel importante na eficiência da rede de sensores, uma vez que permite um ajuste mais adequado do consumo de energia dos sensores que a compõe [Alhajri et al. 2018].

Diversos trabalhos na literatura têm sido propostos com o intuito de se classificar, de forma automática e em tempo real, o tipo de ambiente que circunda um determinado sensor, baseado principalmente na análise das assinaturas do canal de RF utilizado na comunicação com o sensor através da aplicação de sistemas baseados em aprendizagem de máquina [Alhajri et al. 2016], [Zhang et al. 2008], [Alhajri et al. 2018].

Neste sentido, classificadores Fuzzy Auto-Organizáveis (SOF - do inglês *Self-Organising Fuzzy Logic Classifiers*), propostos por [Gu and Angelov 2018] vêm sendo amplamente utilizados para geolocalização, tanto interna quanto externa [Baccar and Bouallegue 2015], [Onofre et al. 2016], [Peña-Rios et al. 2017], [Baccar et al. 2017]. Tendo em vista que dados de RF têm uma alta incerteza associada, levando-se em consideração, que dependendo da intensidade do sinal o objeto pode se encontrar mais próximo de um ambiente que de outro, classificadores fuzzy auto organizáveis representam uma abordagem promissora na solução do problema proposto. Isso porque, além de tratar tal incerteza, propiciam uma adaptabilidade a problemas distintos através da escolha de métricas de distância mais adequadas ao caso em estudo.

Assim, este trabalho propõe uma abordagem baseada em classificadores Fuzzy Auto Organizáveis para a análise da correlação espacial de assinaturas do canal de comunicação RF de redes de sensores sem fio. Possui, desse modo, o intuito de se identificar em tempo real, o tipo de ambiente onde os sensores estão localizados. Além disso, diversas métricas de distância são testadas a fim de se determinar qual destas mais se adequa ao problema proposto. Comparações com outros métodos, presentes na literatura, mostram que o modelo proposto apresenta resultados promissores, tanto em performance quanto em custo computacional.

O restante do artigo está organizado da seguinte forma: a seção 2 trata da

formulação do problema. A seção 3 é dedicada ao modelo SOF. A seção 4 às métricas de distância utilizadas. A seção 5 discute sobre o banco de dados e os resultados das simulações de computador. A seção 6 apresenta as principais conclusões sobre as propostas e trabalhos futuros.

2. Formulação do problema

A classificação do tipo de ambiente onde são alocados os sensores sem fio, utilizados em aplicações de IoT, representam um importante dado no desempenho destas aplicações. Através dessa etapa, é possível determinar um ajuste mais adequado para consumo de energia dos dispositivos envolvidos. Sendo assim, a análise dos coeficientes de correlação espacial dos sinais de RF, presentes no ambiente, podem fornecer um forte indicativo do tipo deste ambiente.

Em particular, em ambientes internos com múltiplos caminhos, como a apresentado na Figura 1, um determinado sinal $Y(f)$ recebido em um ponto de acesso pode conter diversas réplicas de um sinal transmitido $X(f)$. A relação entre $X(f)$ e $Y(f)$ define o que é chamada de Função de Transferência do Canal de comunicação (CTF - do inglês *Channel Transfer Function*), denotada por $H(f)$, a qual contém o efeito dos múltiplos caminhos pelo qual o sinal de RF percorre até ser recebido em um ponto de acesso.



Figura 1. Planta de um ambiente interno de múltiplos caminhos. Fonte: [Alhajri et al. 2018]

Desta forma, $H(f)$ pode ser representada pela superposição dos ganhos associados aos diversos componentes (réplicas) do sinal de RF presentes no ambiente e pode ser expressa por:

$$H(f) = \sum_{l=1}^L a_l \exp[-j(2\pi f\tau - \theta_l)] \quad (1)$$

em que a_l , τ e θ_l são as componentes de amplitude, atraso e fase dos componentes do sinal, respectivamente. L é o número total de componentes dos vários caminhos, f é a frequência do canal classificado. A CTF $H(f)$ é considerada uma assinatura do sinal de RF, sendo única para cada posição no espaço do ambiente.

A autocorrelação complexa da função de transferência do canal de comunicação $H(f)$ é outra métrica de correlação espacial dos sinais de RF. Tal autocorrelação quando no domínio da frequência é conhecida como Função de Coerência de Frequência (FCF - do inglês *Frequency Coherence Function*) e é dada por:

$$R(f) = \int_{-\infty}^{\infty} H(\hat{f})H(\hat{f} + f)d\hat{f} \quad (2)$$

A FCF representa a coerência, no domínio da frequência, do sinal de RF e é interpretada como sendo a autocorrelação da CTF devido a suas diversas mudanças de frequência. Por possuir a característica de mudar lentamente sua forma, quando no domínio do espaço, a CTF se torna uma boa escolha como assinatura de sinal de RF.

3. Classificador Fuzzy Auto-Organizável

O classificador Fuzzy Auto-Organizável (SOF - do inglês *Self-Organising Fuzzy Logic Classifiers*) é um modelo de classificador não-paramétrico baseado em regras que identificam protótipos de dados a partir de uma base de dados observada. Utiliza-se para tanto de um processo de treinamento *offline* que emprega estes protótipos, para então construir um sistema de inferência fuzzy baseado em regras do tipo *AnYa* de ordem 0 [Angelov and Yager 2012] que, por sua vez, é aplicado ao problema de classificação desejado. Proposto em [Gu and Angelov 2018], são utilizadas como métrica para o cálculo da distância, a distância Mahalanobis, Euclideana e Cosseno.

Para fins comparativos, neste artigo, serão aplicadas outras três métricas de distância: Hamming, Minkowski e Manhattan. Visa-se, portanto, ganhos em acurácia na classificação de ambientes internos a partir de medições, em tempo real, do sinal de radiofrequência de uma rede de sensoriamento sem fio.

Uma vez preparado *offline*, no segundo estágio o SOF se encontrado *online*. Haja visto isso, ele é capaz de aprender continuamente a partir do fluxo de dados, de forma a seguir os padrões de mudança nos dados. Dessa forma, ele vai atualizando a estrutura do sistema e os meta-parâmetros recursivamente. Para melhor compreensão, na Figura 2, é apresentada a estrutura do classificador SOF utilizado neste trabalho.

Os meta-parâmetros da abordagem SOF são derivados diretamente dos dados, sendo que ao se alterar o nível de granularidade do modelo é possível se fazer um *trade-off* entre desempenho e eficiência computacional. Assim, o classificador é capaz de lidar com uma ampla variedade de problemas com necessidades específicas.

Além disso, a abordagem SOF permite também a escolha da métricas de distância a ser empregada nos cálculos do algoritmo. A escolha do cálculo de distância utilizado proporciona ainda mais liberdade ao utilizador, tornando a abordagem altamente adaptável a diversas aplicações. Tais métricas são apresentadas a seguir.

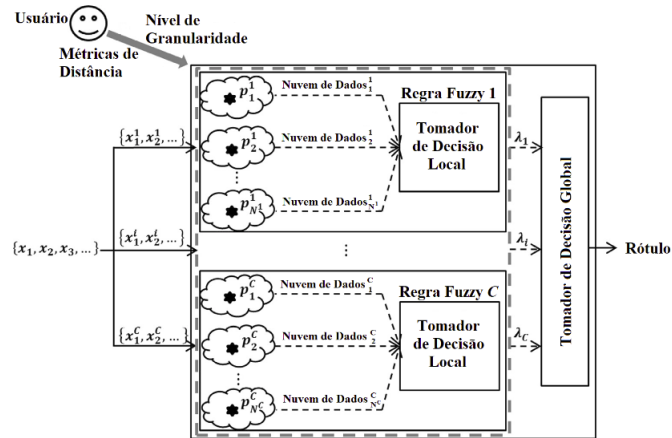


Figura 2. Arquitetura do Classificador Fuzzy Auto-organizável. Fonte: [Gu and Angelov 2018]

Para maiores detalhes a respeito do cálculo das variáveis envolvidas nos dois estágios da abordagem SOF, veja [Gu and Angelov 2018]. O algoritmo e a base de dados utilizada neste artigo estão disponíveis em https://github.com/ualisondias/new_sof.

4. Métricas de Distância

Os autores em [Gu and Angelov 2018] propoem a utilização das distâncias Mahalanobis, Euclideana e Cosseno. Neste artigo, propõe-se variações no modelo em questão, baseadas nas distâncias Hamming, Minkowski e Manhattan visando ganhos de acurácia na classificação de ambientes internos a partir de medições em tempo real do sinal de radiofrequência de uma rede de sensoriamento sem fio.

Às métricas de distâncias entre x_i ; x_j e $\{x\}_K$ adotadas neste trabalho são definidas em [Angelov and Gu 2019].

4.1. Distância Hamming

Utiliza-se a distância Hamming para representar estruturas semânticas complexas com alta fidelidade. É a porcentagem de coordenadas que se diferem, dada por:

$$d_{ham}(x_i, x_j) = \sum_{i,j=1}^n |x_i - x_j| \quad (3)$$

em que n é o número de amostras. Quando $x_i = x_j$, temos $d_{ham} = 0$, enquanto para $x_i \neq x_j$ o valor de $d_{ham} = 1$.

4.2. Distância Minkowski

A distância Minkowski é utilizada quando é necessário identificar e, de fato, ignorar características irrelevantes, e quando existem um número grande de aglomerados anômalos. A métrica é um espaço vetorial normalizado, que pode generalizar a distância Euclideana, a distância Manhattan e a distância Chebychev, dada por:

$$d_{min}(x_i, x_j) = \left(\sum_{l=1}^N (|x_{i,l} - x_{j,l}|)^p \right)^{\frac{1}{p}} \quad (4)$$

em que $0 \leq p \leq 2$.

4.3. Distância Manhattan

Utiliza-se a distância Manhattan em situações que a importância relativa das características deve ser levada em consideração, sendo que ela é uma versão ponderada entre dois vetores normalizados. Às métricas de Manhattan estão todas incluídas na mesma família de funções de distância. Onde essa distância é um caso especial de Minkowski para $p = 1$.

$$d_{man}(x_i, x_j) = \sum_{l=1}^N |x_{i,l} - x_{j,l}| \quad (5)$$

5. Resultados Experimentais

5.1. Base de dados

A base de dados utilizada neste trabalho pode ser encontrada no *website* <https://archive.ics.uci.edu/ml/datasets/2.4+GHZ+Indoor+Channel+Measurements>, e foram publicadas em [Alhajri et al. 2016] e [Alhajri et al. 2018]. Essa base foi produzida no *campus* da Universidade de Khalifa, Emirados Árabes Unidos. Na Figura 1, encontra-se a planta que contém os quatro ambientes analisados, que são o laboratório, salão de esporte, corredor estreito e salão de espera. Os triângulos vermelhos são os transmissores e os asteriscos pretos representam os receptores.

Cada ambiente foi dividido uniformemente com espaços de comprimento de onda ($\lambda = 12,5cm$) e frequência de $f = 2,4Ghz$ selecionadas para examinar as bandas Wi-Fi associadas ao padrão *IEEE 802.11g*. Essas divisões resultaram em um total de 196 posições, sendo que as pequenas variações de escala podem ser melhor capturadas em [Shu et al. 2016]. As medições foram realizadas no ambiente estático, não havendo movimento entre o transmissor e o receptor a fim de não se aumentar o erro associado ao sistema.

O sistema de medição foi constituído por: cabos RF de baixa perda, analisador de rede vetorial (VNA) ZVB14 (usado para medir o coeficiente de transmissão), antenas omnidirecionais de altura igual a 1,5m nas extremidades do receptor e transmissor, sendo que o VNA fazia 10 varreduras consecutivas; cada varredura cobrindo uma faixa de 100 MHz usando 601 pontos de frequência, variando a frequência em 0,167 MHz. O conjunto de dados gerado para cada ambiente consiste em 196 amostras para cada medição, totalizando 1960 amostras.

Desta forma, a base de dados utilizada possui 5 atributos, sendo eles a frequência do sinal, a parte real do parâmetro S_{11} , a parte imaginária do parâmetro S_{11} , a parte real do parâmetro S_{21} e a parte imaginária do parâmetro S_{21} (para maiores informações veja [Alhajri et al. 2018]). Os parâmetros S são as grandezas medidas básicas de um analisador de rede. Eles descrevem como o sistema em

análise, neste caso o ambiente de propagação do sinal de RF, modifica um sinal que é transmitido ou refletido na direção direta ou reversa.

O parâmetro S_{11} é o coeficiente de reflexão da porta 1, ou seja, a relação entre a onda de saída b_1 e a onda incidente a_1 em uma medida direta com a porta correspondente 2: $S_{11} = b_1/a_1$, se $|a_1| > 0$ e $a_2 = 0$. O parâmetro S_{21} é o coeficiente de transmissão para a frente, definido como a razão entre a onda de saída b_2 e a onda incidente a_1 em uma medição à frente com a porta correspondente 2: $S_{21} = b_2/a_1$, se $|a_1| > 0$ e $a_2 = 0$. Na Figura 3, pode-se verificar o fluxo de sinal para uma medição de 2 portas.

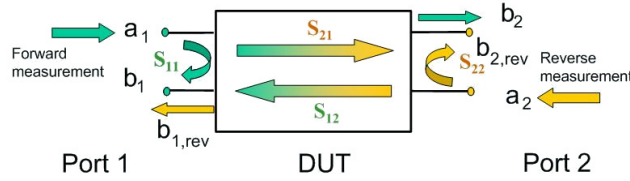


Figura 3. Parâmetros S do analisador de rede vetorial. Fonte: https://www.rohde-schwarz.com/webhelp/zvb_html_usermanual_en/system_overview/measurement_parameters/s-parameters.htm

5.2. Aplicação do SOF

Como mencionado anteriormente, a abordagem SOF possui dois estágios distintos. No primeiro deles, denominado de estágio *offline*, sumarizado no Algoritmo 1, os protótipos dos dados são definidos e utilizados na determinação da base de regras de um modelo *AnYa* estável de ordem 0.

Algoritmo 1 Processo de treinamento *offline* classificador Fuzzy Auto-Organizável

- 1: Calcula D^{MM} em $\{\mathbf{u}\}_{U_K^c}$;
 - 2: Calcula $\mathbf{r}_1 = \arg \max(D_{K^c}^{MM}(\mathbf{u}_i^c))$ e exclui \mathbf{r}_1 de $\{\mathbf{u}\}_{U_K^c}$;
 - 3: $k \leftarrow 1$; $\{\mathbf{r}\} \leftarrow \mathbf{r}_1$; $\{D_{K^c}^{MM}(\mathbf{r})\} \leftarrow D_{K^c}^{MM}(\mathbf{r}_1)$;
 - 4: **while** $U_K^c \neq 0$ **do**
 - 5: * $k \leftarrow k + 1$;
 - 6: * Calcula $\mathbf{r}_k = \arg \min(d(\mathbf{r}_{k-1}, \mathbf{u}_i^c))$ e exclui \mathbf{r}_k de $\{\mathbf{u}\}_{U_K^c}$;
 - 7: * $\{\mathbf{r}\} \leftarrow \{\mathbf{r}\} + \mathbf{r}_k$; $\{D_{K^c}^{MM}(\mathbf{r})\} \leftarrow \{D_{K^c}^{MM}(\mathbf{r})\} + D_{K^c}^{MM}(\mathbf{r}_k)$;
 - 8: **end while**
 - 9: Identifica $\{\mathbf{p}_0\}$;
 - 10: Organiza nuvens de dados vizinhos $\{\mathbf{p}\}_0$;
 - 11: Identifica $\{\varphi\}_0$ a partir da nuvem de dados;
 - 12: Calcula D^{MM} em $\{\varphi\}_0$;
 - 13: Identifica $\{\varphi\}^{vizinho}$;
 - 14: Identifica $\{\mathbf{p}\}^c$;
 - 15: Cria a c^{th} regra *fuzzy* com $\{\mathbf{p}\}^c$;
-

em que $\{\mathbf{u}\}_{U_K^c}$ é a base de dados utilizada, sem amostras repetidas, D^{MM} é a densidade multimodal dos dados, \mathbf{r} é uma lista em que as amostras são classificadas de

acordo com sua respectiva densidade multimodal, \mathbf{p}_0 representam os protótipos dos dados, $\{\varphi\}_0$ são os centros das nuvens de dados, $\{\varphi\}^{vizinho}$ são os centros das nuvens de dados vizinhos à nuvem i e $\{\mathbf{p}\}^c$ é o protótipo mais significativo da c^{th} classe.

Já no segundo estágio da abordagem SOF, denominado de estágio *online*, as regras identificadas são atualizadas de acordo com os dados de fluxo contínuo, acompanhando assim as possíveis alterações nestes dados. Este estágio é sumarizado no Algoritmo 2, em que $\mu_{k^c}^c$, $X_{k^c}^c$ e $\Sigma_{k^c}^c$ são metaparâmetros do modelo, $\mathbf{G}_{k^c}^{c,L}$ é o raio médio das áreas locais de influência, N^c é o número de nuvens de dados e $\mathbf{S}_{N^c}^c$ é o suporte (número de membros) da nuvem de dados.

Algoritmo 2 Processo de treinamento *online* classificador Fuzzy Auto-Organizável

```

1: while uma nova amostra de dado da  $c^{th}$  classe  $x_{k^c+1}^c$  está disponível (ou até
   interrupção) do
2:   Atualiza  $\mu_{k^c}^c$ ,  $X_{k^c}^c$ ,  $\Sigma_{k^c}^c$ ,  $\mathbf{G}_{k^c}^{c,L}$  para  $\mu_{k^c+1}^c$ ,  $X_{k^c+1}^c$ ,  $\Sigma_{k^c+1}^c$ ,  $\mathbf{G}_{k^c+1}^{c,L}$ ;
3:   Calcula  $D$  em  $x_{k^c+1}^c$  e  $\{\mathbf{p}\}^c$ ;
4:   if um novo protótipo for adicionado a regra) ou (um novo protótipo e a nova
   nuvem for adicionado a regra) then
5:     *  $N^c \leftarrow N^c + 1$ ;  $\mathbf{p}_{N^c}^c \leftarrow x_{k^c+1}^c$ ;  $\mathbf{S}_{N^c}^c \leftarrow 1$ ;  $\{\mathbf{p}\}^c \leftarrow \{\mathbf{p}\}^c + \mathbf{p}_{N^c}^c$ ;
6:   else
7:     * Encontra  $\mathbf{p}_{n^*}^c$ ;
8:     *  $\mathbf{p}_{n^*}^c \leftarrow \frac{\mathbf{S}_{n^*}^c}{\mathbf{S}_{n^*}^c + 1} \mathbf{p}_{n^*}^c + \frac{1}{\mathbf{S}_{n^*}^c + 1} x_{k^c+1}^c$ ;  $\mathbf{S}_{n^*}^c \leftarrow \mathbf{S}_{n^*}^c + 1$ ;
9:   end if
10:   $K^c \leftarrow K^c + 1$ ;
11:  Atualiza as regras fuzzy
12: end while

```

Para maiores detalhes a respeito do cálculo das variáveis envolvidas nos dois estágios da abordagem SOF, veja [Gu and Angelov 2018]. O algoritmo e a base de dados utilizada neste artigo estão disponíveis em https://github.com/ualisondias/new_sof.

Assim, o coeficiente de relação espacial, dado por $C(n)$ em [Zhang et al. 2008], que pode ser visto como a similaridade entre as assinaturas dos sinais de RF pode ser calculado como:

$$C(n) = \frac{\langle S_r, S^{(n)} \rangle}{\|S_r\| \|S^{(n)}\|} \quad (6)$$

em que S_r é a assinatura de RF correspondente ao CTF ou FCF em uma posição de referência específica, enquanto $S^{(n)}$ é a assinatura RF na posição n .

Quando os valores de $C(n)$ são pequenos, significa que existe uma diferença significativa entre os valores da assinatura RF na posição n e os valores de referência; quando os valores são altos, indica alta similaridade entre as assinaturas RF na posição n e os valores de referência RF. A forma dos sinais de correlação espacial da a CTF e da FCF para os quatro ambientes analisados são apresentados nas Figuras 4 e 5 respectivamente.

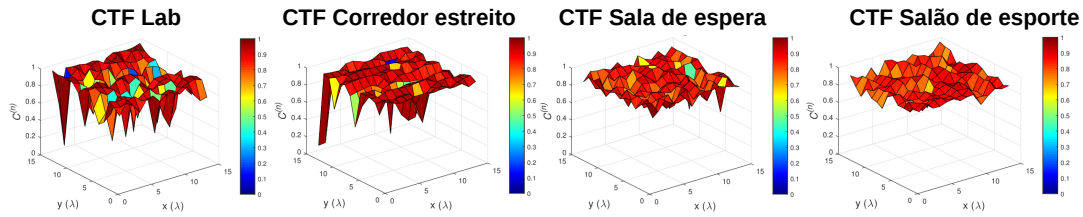


Figura 4. Correlação espacial baseada na CTF para os diversos ambientes analisados.
Fonte: [Alhajri et al. 2018]

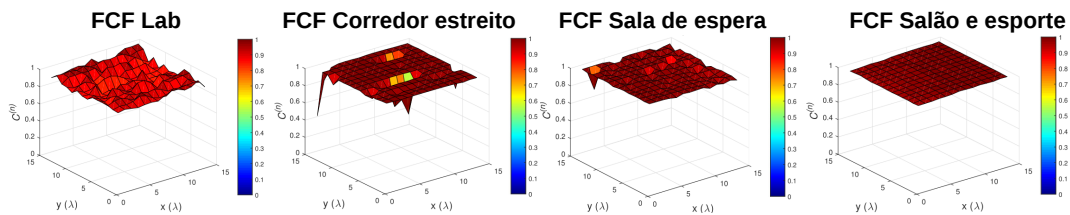


Figura 5. Correlação espacial baseada na FCF para os diversos ambientes analisados.
Fonte: [Alhajri et al. 2018]

Foram obtidos aleatoriamente 100.000 amostras, sendo 50.000 para teste, 7.500 para treinamento *offline* e 42.500 de treinamento *online*. Essa divisão totaliza 50% de treinamento e 50% teste, sendo que para treinamento destina-se 15% para treinamento *offline* e 35% para treinamento *online*.

5.3. Análise das métricas

Neste trabalho as diversas métricas de distância apresentadas foram testadas a fim de se verificar qual delas mais se adaptaria ao problema proposto. Também foram testados diversos níveis de granularidade da entrada a fim de se determinar qual seria o melhor balanço entre desempenho e custo computacional. A Figura 6 apresenta uma análise de precisão em função da variação da granularidade para cada distância distinta a partir da base de dados descrita na Seção 5.1.

Analisando a Figura 6, é possível observar que a escolha da granularidade dos dados de entrada tem alto impacto nos valores de precisão do modelo para determinadas métricas de distância, enquanto que para outras este impacto é quase inexistente. Por exemplo, para as distância Cosseno e Hamming, não existe alteração de *Acc* com o aumento da granularidade, mas o mesmo não é observado para as distâncias Euclidean, Mahalanobis, Manhattan e Minkowski.

Na Tabela 1, são apresentados os resultados obtidos pela aplicação da abordagem SOF ao problema de classificação de ambientes com base na sinal de correlação espacial de sinais de RF para as diversas métricas de distância discutidas. Nela é possível verificar que quando utilizada a distância Hamming a precisão do modelo atinge 73,15%, sendo o menor valor encontrado. Já para a distância Manhattan, que é uma das variações da distância Minkowski tendo $p = 1$, obteve-se 98,22% , sendo

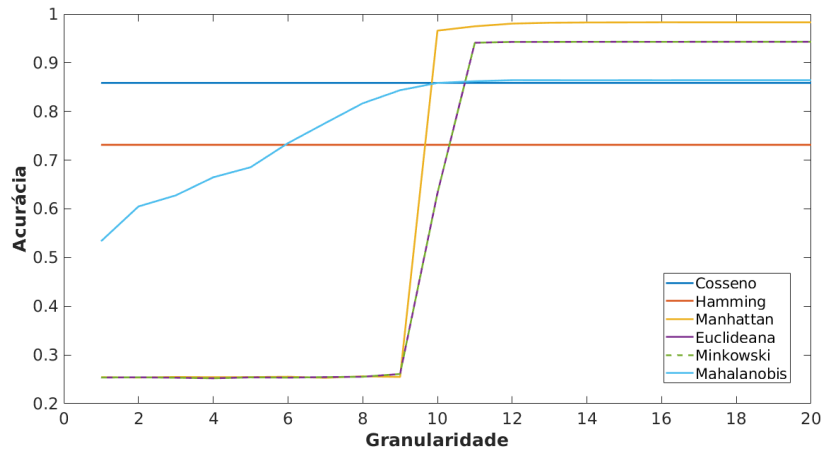


Figura 6. Precisão em função da granularidade para as distância apresentadas neste artigo

este o melhor resultado dentre todos. As distâncias Cosseno e Mahalanobis obtiveram resultados bem próximos, assim como as distâncias Euclideana e Minkowski.

Tabela 1. Resultados Obtidos.

Abordagem	Acurácia (%)	Tempo (s)
SOF- Euclideana	94,11	45.18
SOF- Mahalanobis	86,42	155.42
SOF- Cosseno	85,86	60.88
SOF- Hamming	73,15	47.40
SOF- Manhattan	98,22	57.43
SOF- Minkowski	94,28	47.13

É possível avaliar também os valores de tempo de processamento, visto que as especificações do computador usado para executar o algoritmo e gerar os modelos de aprendizado são fornecidas da seguinte forma: CPU Intel Core i7-5500U (2 cores and 4 Threads de 2.4GHz e memória cache de 4MB), memória RAM de 8GB e sistema operacional Linux Fedora 29. O melhor resultado de tempo foi obtido pela aplicação da distância Euclideana, enquanto que o pior resultado foi obtido pela distância de Mahalanobis. O tempo de processamento para as distâncias Cosseno e Manhattan ficaram bem próximos, assim como, Hamming e Minkowski.

Na Figura 7 é apresentada a matriz de confusão para o passo de teste, mostrando as classes de saída para os quatro ambientes diferentes denotados pela primeira linha/coluna (corredor estreito), segunda linha/coluna (laboratório), terceira linha/coluna (sala de espera) e quarta linha/coluna (salão de esporte), em função das classes alvo na vertical. Nela é possível visualizar que a métrica com melhor desempenho dentre as avaliadas, a distância Manhattan, que obteve 98,22% de precisão para uma granularidade igual a 13.

	100%	0.0%	0.0%	0.0%
	12540	0	0	0
Classes de Saída	0.0%	96.35%	1.22%	2.41%
	0	12083	154	303
	0.0%	0.68%	97.79%	1.52%
	0	86	12263	191
	0.0%	0.71%	0,54%	98.74%
	0	90	68	12382
	Classes Alvo			

Figura 7. Matriz de confusão utilizando distância Manhattan

6. Conclusão

Neste trabalho foi apresentada uma abordagem, baseada em classificadores fuzzy auto-organizáveis aplicados à análise de dados de correlação espacial de sinais de RF em redes de sensores sem fio, com o intuito de se classificar o ambiente no qual estes sensores estão localizados. O classificador fuzzy se mostrou livre de parâmetros pré-definidos ou suposições anteriores, sendo impulsionado apenas pelos dados observados empiricamente, identificando protótipos dos dados de treinamento *offline* e aprendendo continuamente com os dados de fluxo de forma recursiva. Diversas métricas de distância e vários níveis de granularidade foram utilizados, sendo que os resultados experimentais mostraram que o método proposto apresentou um alto desempenho e baixo custo computacional quando aplicado a uma base de dados composto por medições, em tempo real do sinal RF, em ambientes internos com múltiplos caminhos, mesmo esta sendo uma base com grande densidade de dados.

Referências

- [Alhajri et al. 2016] Alhajri, M., Alsindi, N., Ali, N., and Shubair, R. (2016). Classification of indoor environments based on spatial correlation of rf channel fingerprints. *Antennas and Propagation, 2016 IEEE International Symposium on*, pages 1447–1448.
- [Alhajri et al. 2018] Alhajri, M., Raed, M., Ali, N., and Shubair, R. (2018). Classification of indoor environments for iot applications: A machine learning approach. *IEEE Antennas and Wireless Propagation Letters*, 17(12):2164–2168.
- [Angelov and Yager 2012] Angelov, P. and Yager, R. (2012). A new type of simplified fuzzy rule-based system. *International Journal of General Systems*, 41(2):163–185.
- [Angelov and Gu 2019] Angelov, P. P. and Gu, X. (2019). *Empirical approach to machine learning*. Springer, 1st 2019 edition.
- [Baccar and Bouallegue 2015] Baccar, N. and Bouallegue, R. (2015). Intelligent type 2 fuzzy-based mobile application for indoor geolocation. *2015 23rd International Conference on Software, Telecommunications and Computer Networks*, pages 165–169.

- [Baccar et al. 2017] Baccar, N., Jridi, M., and Bouallegue, R. (2017). Adaptive neuro-fuzzy location indicator in wireless sensor networks. *Wireless Personal Communications*, 97(2):3165–3181.
- [Gu and Angelov 2018] Gu, X. and Angelov, P. P. (2018). Self-organising fuzzy logic classifier. *Information Sciences*, 447:36–51.
- [Onofre et al. 2016] Onofre, S., Caseiro, B., Pimentão, J. P., and Sousa, P. (2016). Using fuzzy logic to improve ble indoor positioning system. *Doctoral Conference on Computing, Electrical and Industrial Systems*, 470:169–177.
- [Peña-Rios et al. 2017] Peña-Rios, A., Hagra, H., Gardner, M., and Owusu, G. (2017). A type-2 fuzzy logic based system for asset geolocation within augmented reality environments. *2017 IEEE International Conference on Fuzzy Systems*, pages 1–6.
- [Shu et al. 2016] Shu, Y., Huang, Y., Zhang, J., Coué, P., Cheng, P., Chen, J., and Shin, K. G. (2016). Gradient-based fingerprinting for indoor localization and tracking. *IEEE Transactions on Industrial Electronics*, 63(4):2424–2433.
- [Tiwari et al. 2015] Tiwari, P., P. Saxena, V., Mishra, R. G., and Bhavsar, D. (2015). Wireless sensor networks: Introduction, advantages, applications and research challenges. *Open International Journal of Technology Innovations and Research*, 14.
- [Zhang et al. 2008] Zhang, J., Firooz, M. H., Patwari, N., and Kasera, S. K. (2008). Advancing wireless link signatures for location distinction. *Proceedings of the 14th ACM international conference on Mobile computing and networking*, pages 26–37.