Distância invariante à complexidade baseada em dimensão fractal para classificação de séries temporais

Ronaldo C. Prati[1] **e Gustavo E. A. P. A. Batista**[2]

[1]Centro de Matemática, Computação e Cognição – Universidade Federal do ABC (UFABC)
Rua Santa Adélia, 166 – 09.210-170, Bangú – Santo André, SP – Brasil

[2]Instituto de Ciências Matemáticas e de Computação – Universidade de São Paulo
Avenida Trabalhador são-carlense, 400 – 13.566-590, Centro – São Carlos, SP – Brasil

ronaldo.prati@ufabc.edu.br, gbatista@icmc.usp.br

Abstract. *This paper proposes the use of the complexity given by fractal dimension to develop a complexity invariant measure for time series classification. Fractal dimension has a broad application as it is independent from the series amplitude. Empirical evidence shows that fractal dimension is competitive to the original measure.*

Resumo. *Neste artigo propomos o uso da complexidade medida por meio da dimensão fractal como fator de correção para uma medida de distância invariante à complexidade para a classificação de séries temporais. Dimensão fractal é uma abordagem mais geral do que a originalmente proposta, pois não depende da amplitude da série. Resultados experimentais mostram que dimensão fractal é competitiva com a medida originalmente proposta.*

1. Introdução

Séries temporais é um importante tópico de pesquisa em diferentes áreas de aplicação, tais como em medicina com exames como eletroencefalograma e eletrocardiograma; no mercado financeiro com a previsão do preço de *commodities* e valor de ações; em processos dinâmicos na academia com a análise de sinais captados por rádiotelescópios; e na indústria, agricultura e pecuária com redes de sensores (para a agriculta de precisão, previsões de variações climáticas, etc), dentre muitas outras aplicações.

Devido a relevância de séries temporais, a comunidade científica tem proposto nas últimas décadas um grande número de técnicas para tarefas como classificação, agrupamento, previsão, detecção de anomalias, entre muitas outras. Diversos trabalhos apontam evidências empíricas que, para séries temporais, métodos baseados em distâncias provêem resultados muito competitivos, frequentemente superioes a outras abordagens mais complexas. Por exemplo, em classificação, uma estratégia simples de 1-vizinho mais próximo (1NN) fornece resultados difíceis de serem superados [Ding et al. 2008]; em detecção de anomalias [Chandola et al. 2009] indica que, após uma extensiva avaliação empírica, os métodos baseados em distâncias fornecem o melhor desempenho geral entre todas as técnicas avaliadas; e em agrupamento de dados, trabalhos recentes tem sugerido que a escolha do algoritmo de agrupamento é muito menos importante do que a escolha da medida utilizada, e que a distância *Dynamic Time Warping* (DTW) provê resultados superioes a outras medidas [Zhu et al. 2012].

O desempenho de uma medida de distância em uma tarefa de mineração de dados está associada à capacidade dessa medida em capturar corretamente as "invariância" requeridas pelo domínio de aplicação [Batista et al. 2011]. Um exemplo é a invariância à amplitude, de forma que dados em diferentes unidades (tais como temperaturas Celsius e Fahrenheit) possam ser diretamente comparados. Outras formas de invariância incluem invariância à fase (utilizada com dados periódicos), escala local (*warping*) e uniforme [Keogh 2003] e oclusão.

Recentemente [Batista et al. 2011] propôs uma nova forma de invariância denominada invariância à complexidade. Foi observado que diversos problemas apresentam séries temporais com diferentes complexidades e que pares de objetos complexos, mesmo aqueles que são visualmente similares, tendem a ser considerados, pelas medidas de distância atuais, mais distantes do que pares de objetos simples. Os autores propuseram uma medida de distância corrigida, denominada CID (distância invarante à complexidade, do inglês *Complexity-Invariant Distance*).

Os resultados reportados em [Batista et al. 2011] mostram a validade do uso da CID em diversos problemas de classificação de séries temporais. Entretanto, como os autores discutem em seu artigo, existem muitas formas de estimar a complexidade de uma série temporal, e o fator de correção proposto no artigo faz uso de uma das possíveis estimativas, a qual foi escolhida empiricamente.

Neste artigo, investigamos o uso da dimensão fractal como estimativa do fator de correção para uma medida de complexidade invariante à distância. A dimensão fractral é uma das medidas de complexidade mais utilizadas na literatura, pois permite estimar a dimensão intrínseca dos dados. As contribuições desse artigo são:

- Validação experimental, utilizando uma abordagem diferente da proposta em [Batista et al. 2011], de que invariância à complexidade é um fator relevante na classificação de séries temporais;
- Avaliação experimental que mostra que o uso de dimensão fractal é uma abordagem competitiva ao proposto originalmente pelos autores;
- Uma alternativa para a correção de complexidade que é independente do tamanho da série. O fator de correção originalmente proposto na CID leva em consideração o tamanho da série "linearizada", de maneira que somente é aplicável a séries de mesmo tamanho. Dimensão fractal não tem essa limitação, o que permite que ela seja aplicada na classificação séries de diferentes tamanhos.

2. Revisão das invariâncias conhecidas

Frequentemente dados temporais possuem distorções indesejáveis. Tais distorções podem fazer com que medidas de distância não consigam capturar adequadamente a similaridade entre as séries temporais, associando distâncias demasiadamente grandes a objetos similares. Nesta seção são revisadas todas as distorções conhecidas em dados temporais, bem como os métodos utilizados para se obter invariância a essas distorções.

Invariância à amplitude e *offset* Duas séries temporais medidas em escalas diferentes, como temperaturas em Celsius e Fahrenheit, não serão consideradas similares mesmo se possuírem uma forma similar. Para medir a verdadeira similaridade é necessário fazer

com que as amplitudes sejam as mesmas. Por exemplo, as séries temporais da Figura 1-acima possuem um formato similar, mas possuem uma grande distância euclidiana devido à diferença de amplitudes. De forma similar, mesmo se duas séries temporais possuem amplitudes idênticas, elas podem ter *offsets* diferentes (diferentes valores médios). Como mostrado na Figura 1-abaixo, mesmo uma pequena mudança de *offset* pode rapidamente dominar a distância euclidiana, levando a falsos negativos, por exemplo, batidas cardíacas não detectadas. Ambas invariâncias à amplitude e *offset* podem ser obtidas trivialmente por meio da normalização em z-*scores* dos dados [Faloutsos et al. 1994].

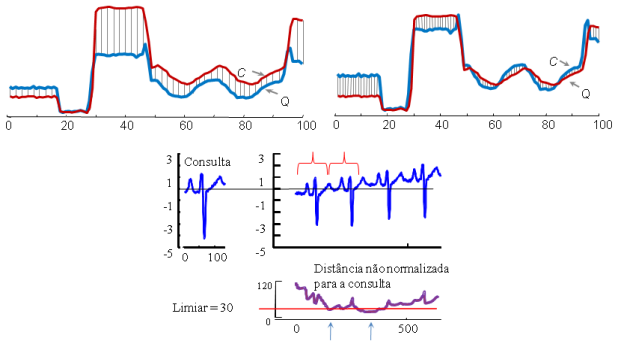

Figura 1. acima) Se comparadas antes da normalização de amplitude, essas duas séries temporais aparentam ser bem diferentes. abaixo) Quando um objeto de consulta de batimento cardíaco é comparado com uma sequência de batimentos, os dois primeiros batimentos casam bem, mas a mudança de *offset* faz com que os batimentos subsequentes não sejam detectados

Invariância à escala local (“*warping*”) Essa invariância é necessária para quase todos os sinais de origem biológica, incluindo a captura de movimentos, reconhecimento de escrita e eletrocardiogramas. Na Figura 2 é mostrado um exemplo de uma série temporal representando os comportamentos de dois insetos que somente casam quando a invariância à escala local é utilizada. Dada a ubiquidade dos domínios que requerem essa invariância, existem centenas de artigos sobre esse tópico. Entretanto, evidências empíricas recentes sugerem que uma técnica proposta há quarenta e cinco anos, *Dynamic Time Warping* (DTW), funciona excepcionalmente bem [Ding et al. 2008].

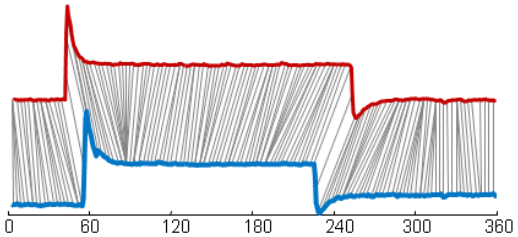

Figura 2. Duas séries temporais representando comportamentos de insetos casam bem quando há invariância à escala local. O alinhamento foi calculado com o algoritmo DTW

Invariância à escala uniforme Em contraste à escala local realizada pela distância DTW, muitos conjuntos de dados requerem um redimensionamento global. Por exemplo,

na Figura 3 são mostradas duas séries temporais de expressão gênica de levedura de dois genes conhecidamente relacionados [Li et al. 2002]. O alinhamento da sequência mais curta com o prefixo da sequência mais longa resulta em um casamento ruim. Entretanto, se a sequência mais curta for globalmente reescalada por um fator de 1.41, resulta em um casamento melhor. A maior dificuldade em se criar uma invariância à escala uniforme é que tipicamente não se conhece o fator de escala a priori, restando a possibilidade de testar todas as possibilidades dentro de um determinado intervalo [Keogh 2003].

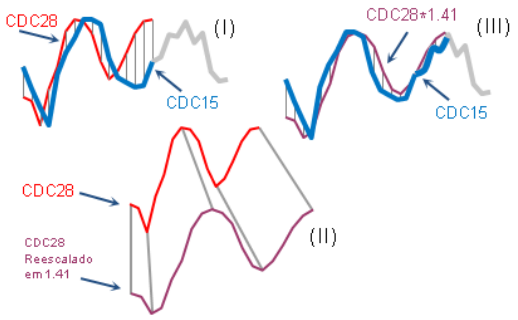

Figura 3. (I) A expressão completa de um gene, CDC28, casa mal com o prefixo de um gene correlacionado, CDC15. II) Se a série temporal for reescalada por um fator de 1.41, ela se torna mais similar ao gene CDC15 (III)

Invariância à fase Essa forma de invariância é importante para casar séries temporais periódicas tais como curvas de luz estrelar [Keogh et al. 2009], batimentos cardíacos, etc. Ela também é importante para casar formas bidimensionais que foram convertidas para séries "temporais" unidimensionais por meio de um truque de representação que tem ganhado popularidade nos últimos anos [Keogh et al. 2009]. Diversos autores têm sugerido obter essa invariância por meio do uso de alinhamentos cardinais para os quais todas as séries temporais são alinhadas. Entretanto, evidências recentes sugerem que essa abordagem pode não fornecer bons resultados [Zunic et al. 2006], e a única forma atualmente conhecida para garantir invariância à fase é testar todos os possíveis alinhamentos, como é mostrado na Figura 4.

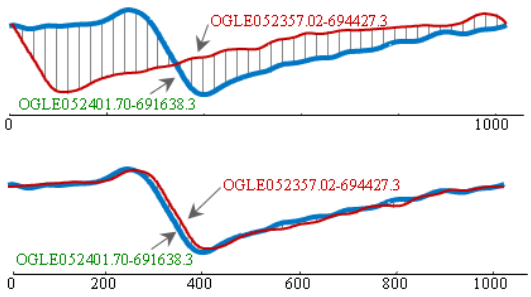

Figura 4. acima) Duas curvas de luz estrelar estão obviamente fora de fase. abaixo) Mantendo uma série temporal em posição fixa, e testando todos os deslocalmentos circulares da outra, é possível obter invariância à fase

Invariância à oclusão Essa forma de invariância ocorre em domínios nos quais as séries temporais podem ter uma pequena subsequência faltante. Um exemplo visual é apresen-

tado na Figura 7, na qual uma imagem de um crânio antigo casa quase perfeitamente com uma imagem de um crânio moderno, apesar do crânio antigo ter a região do nariz faltante. Repare que foi realizada uma transformação de representação entre uma imagem de crânio bidimensional e uma série "temporal" unidimensional.

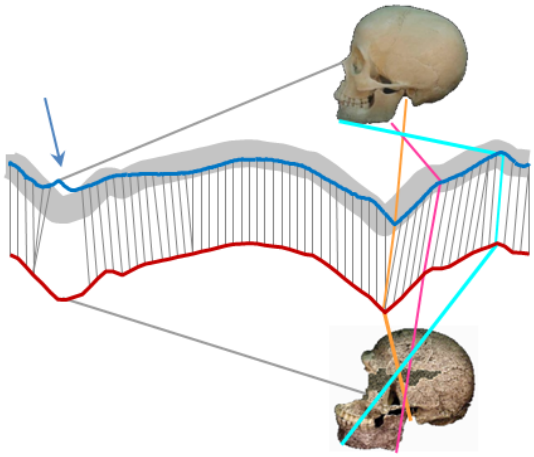

Figura 5. Invariância à oclusão pode ser obtida seletivamente recusando-se a casar subseções de uma série temporal. Neste exemplo, a distância se torna robusta a falta da região do nariz no crânio antigo

Invariância à complexidade Invariância à complexidade utiliza informação sobre diferenças entre as complexidades dos objetos a serem comparados como um fator de correção para as medidas de distância existentes. Por exemplo, a distância euclidiana, $ED(Q, C)$, entre duas séries temporais Q e C, pode se tornar invariante à complexidade por meio da introdução do seguinte fator de correção:

$$CID(Q, C) \equiv ED(Q, C) \times CF(Q, C) \tag{1}$$

Na qual CF é um fator de correção de complexidade definido como:

$$CF(Q, C) \equiv \frac{\max(CE(Q), CE(C))}{\min(CE(Q), CE(C))} \tag{2}$$

e $CE(T)$ é a estimativa de complexidade da série temporal T. Como discutido anteriormente, CE pode ser estimado de diversas maneiras. O fator de correção CF visa contabilizar diferenças entre as complexidades das séries temporais comparadas. CF faz com que séries temporais com grandes diferenças de complexidade se tornem mais distantes. Caso todas as séries temporais tenham a mesma complexidade, CID simplesmente degenera para a distância euclidiana.

Originalmente, CID utiliza uma estimativa de complexidade bastante simples. Ela é baseada na intuição de que se nós pudéssemos "esticar" uma série temporal até que ela se torne um segmento de reta, uma série temporal complexa resultaria em um segmento de reta mais longo do que uma série temporal simples. Na Figura 6, essa ideia é ilustrada com alguns exemplos. Essa estimativa de complexidade pode ser calculada por

meio da seguinte equação:

$$CE(Q) \equiv \sqrt{\sum_{i=1}^{n-1}(q_i - q_{i+1})^2} \tag{3}$$

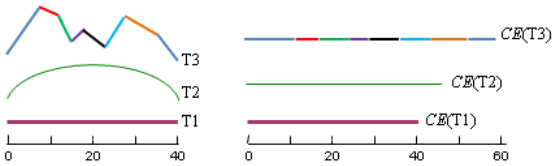

Figura 6. esquerda) Três séries temporais podem ter as suas complexidades medidas "esticando" cada uma delas e medindo o comprimento dos segmentos de reta resultantes (direita).

Em uma análise da Equação 3 fica evidente a simplicidade da medida de complexidade, bem como das suas limitações. Essa medida de complexidade deve ser aplicada para séries com o mesmo número de observações, uma vez que somente as diferenças entre os valores das observações são levadas em consideração, ignorando-se as diferenças de tempo em que as observações ocorreram. Essa limitação não restringe o uso da complexidade quando aliada a uma medida de distância como a distância euclidiana, pois essa medida também requer que as séries possuam o mesmo número de observações. Mas restringe o uso de outras distâncias mais flexíveis, como no caso da distância DTW. Ainda, a medida definida pela Equação 3 requer que as séries a serem comparadas estejam previamente normalizadas em amplitude e *offset*. Embora esse seja uma normalização padrão para a grande maioria dos domínios de aplicação, uma medida de complexidade que não requeira tal normalização pode ser bastante útil em outros domínios nos quais diferenças de escala e *offset* podem caracterizar atributos legítimos para a classificação.

3. Dimensão fractal de séries temporais

A dimensão fractal usa conceitos de auto-similaridade para estimar a rispidez (ou suavidade) de uma série temporal. Essa característica a torna uma medida possivelmente interessante para estimar a complexidade de uma série. Existem diferentes maneiras de calcular a dimensão fractal de uma série, apesar da maiora dos métodos geralmente seguir um esquema comum [Gneiting et al. 2010]:

- Alguma proprieda numérica ρ, que depende de alguma escala ϵ, é calculada a partir da série (no domínio temporal ou da frequência);
- Uma lei de potência assintótica $\rho \propto b$ é derivada ou postulada;
- O parâmetro b da lei de potência é uma função linear da dimensão fractal FD;
- Dessa maneira, FD pode ser estimada usando regressão linear de $\log \rho(\epsilon)$ sobre $\log \epsilon$, variando-se o a escala ϵ, com ênfase nos menores valores observados de ϵ.

Neste trabalho, investigamos as seguintes maneiras de se calcular a dimensão fractal, implementadas no pacote fractaldim [Gneiting et al. 2010], do software R:

Box counting A ideia básica de estimar FD usando box counting é simples: inicialmente, o gráfico da série é enquadrado em uma caixa simples. A caixa é então dividida em quatro quadrantes, e o número de células necessárias para cobrir o gráfico da série é contado. O próximo passo consiste em dividir cada quadrante em quatro subquadrantes, e o processo continua até cada caixa tenha a mesma resolução dos dados, contando-se a fração de quadrantes necessária para cobrir todo o gráfico em cada passo. Se $\rho(\epsilon)$ denota a fração de caixas necessárias com a largura ϵ de cada caixa naquele passo, o estimador box-count é a inclinação da reta obtida pela regressão linear de $\log \rho(\epsilon)$ sobre $\log \epsilon$.

Hall-Hood é uma variação do box-counting no qual ao invés de calcular a fração de caixas em cada passo, considera a área das caixas que cobrem o gráfico da curva para estimar FD.

Estimadores baseados em variância usam a variância da série a cada escala ϵ como medida de auto-similaridade. A variância pode ser calculada como:

$$var(Q) = \sum_{i}^{\rho(\epsilon)} (q_i - q_{i+\epsilon})^p \tag{4}$$

O gráfico de $\log var(Q)$ versus $\log \epsilon$ fornece o the semi-variograma de uma série temporal, e FD pode ser computada com base na inclinação do semi-variograma. **Variação** usa o parâmetro p definido pelo usuário. Se a potência p em 4 for igual a 2, 1, $1/2$ temos o **Variograma**, **Madograma** e **Rodograma**, respectivamente. **Incr1** usa uma fórmula alternativa para o calculo da variância, basea em derivadas de segunda ordem (para mais detalhes, veja [Gneiting et al. 2010])

Estimadores espectrais e wavelets operam no domínio da frequência ao invés do domínio temporal. **Wavelets** usa os coeficientes da transformada de wavelets, enquanto **Periodograma** e **DCT-II** são baseados na densidade espectral da série. De maneira análoga ao semi-variograma, é computado o semi-periodograma e a dimensão fractal é calculada a partir da inclinação da reta ajustada por regressão de diferentes frequência com a densidade espectral (para mais detalhes, veja [Gneiting et al. 2010]).

4. Resultados experimentais

Nesta seção, descrevemos os resultados experimentais obtidos a partir do uso da dimensão fractal, calculados pelos métodos descritos na Seção 3, para calcular a complexidade de uma série temporal ao invés da medida originalmente proposta (Equação 3). Para a realização dos experimentos, seguimos os mesmos procedimentos experimentais adotados em [Batista et al. 2011], utilizando todos os conjuntos de dados para classificação de séries temporais disponíveis no repositório [Keogh et al. 2011].

No total, a avaliação incluiu 43 conjuntos de dados de diferentes domínios, incluindo medicina, entomologia, engenharia, astronomia, processamento de sinais, entre outros. Uma descrição detalhada das características dos conjuntos de dados pode ser encontrada na página do repositório, e não é reproduzida neste artigo por questões de espaço.

Os experimentos foram executados utilizando o software R, e para o cálculo da distância DTW foi utilizado o pacote dtw [Giorgino 2009].

A acurácia do classificador 1NN foi avaliada utilizando as mesmas partições de treinamento e teste disponibilizadas no repositório. Avaliamos a acurácia utilizando distância euclidiana, a distância CID originalmente proposta (descrita na Seção 2), a qual chamaremos CID_{ori} e as variações de CID utilizando a dimensão fractal, a qual chamaremos de CID_{df}.

Na Figura 7 estão sumarizados os resultados da comparação de CID_{ori} e CID_{df}. Nessa figura, o gráfico da esquerda é uma reprodução dos resultados obtidos em [Batista et al. 2011], que comparam CID_{ori} com a distância euclidiana. Já o gráfico da direita apresenta a comparação do classificador com maior acurácia dentre aqueles que usam a dimensão fractal para correção de complexidade CID_{df} com a distância euclidiana. Como pode ser observado nessa figura, para a maioria dos conjuntos de dados utilizados, tanto CID_{ori} quanto CID_{df} apresentam uma melhor acurácia do que a distância euclidiana apenas. Entretanto, enquanto para CID_{ori} em oito dos 43 conjuntos de dados a acurácia foi menor do que a distância euclidiana apenas, e houve três empates, para CID_{df} houve apenas um conjunto de dados em que a acurácia foi ligeiramente menor e dois empates.

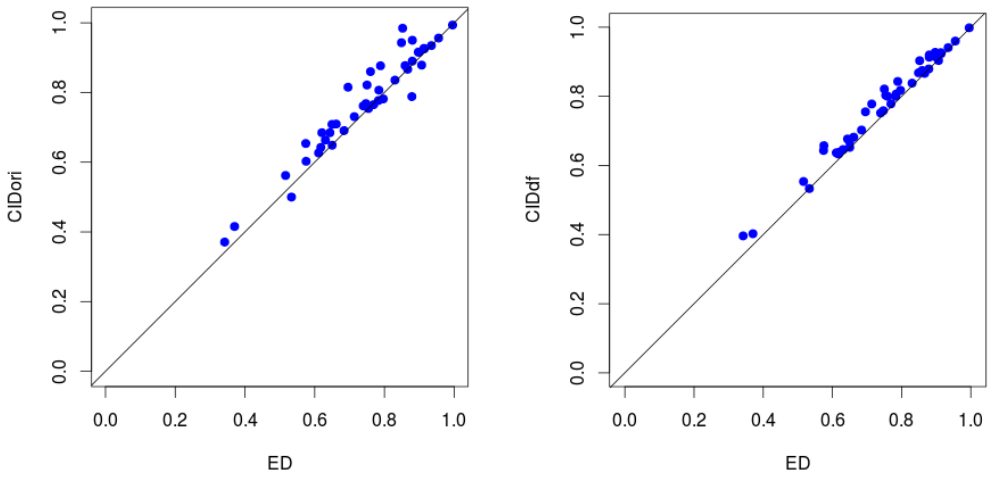

Figura 7. esquerda) comparação de desempenho entre CID_{ori} e a distância euclidiana; direita) comparação de desempenho entre CID_{df} e a distância euclidiana. Pontos acima da diagonal indicam que CID é melhor que ED

Analisando esses resultados podemos verificar que, usando uma abordagem diferente da proposta em [Batista et al. 2011], que a invariância de complexidade é um fator relevante na classificação de séries temporais. Além disso, o fato de em apenas um dos conjuntos de dados utilizando CID_{df} a acurácia foi ligeiramente menor do que o classificador obtido utilizando apenas distância euclidiana indica que há espaço para melhora a partir de CID_{ori}.

Na Figura 8 é apresentada uma comparação entre CID_{ori} e CID_{df}. Essa figura

mostra que CID_{df} é competitivo com CID_{ori}. Em 20 conjuntos de dados, CID_{df} obteve uma acurácia melhor que CID_{ori}, houve quatro empates e em 19, CID_{ori} obteve uma acurácia superior.

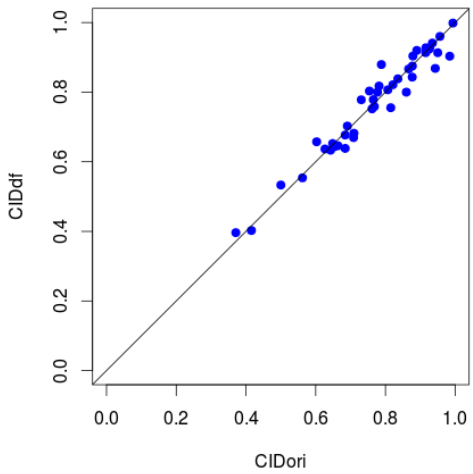

Figura 8. Comparação de desempenho entre CID_{ori} e CID_{df}. Pontos acima da diagonal indicam que CID_df é melhor que CID_{ori}

De maneira análoga à Figura 7, na Figura 9 é apresentada uma comparação da acurácia obtida por meio da distância DTW com a versão invariante à complexidade $CID(DTW)_{ori}$ da distância DTW originalmente proposta, além da variação utilizando dimensão fractal para o cálcula da complexidade $CID(DTW)_{df}$.

Essa figura apresenta um padrão semelhante ao apresentado na Figura 7. No gráfico a esquerda, é mostrada uma reprodução dos resultados originais obtidos em [Batista et al. 2011], que compara $CID(DTW)_{ori}$ com DTW, e no gráfico da direita apresenta a comparação do classificador com maior acurárcia dentre aqueles que usam a dimensão fractal para correção de complexidade $CID(DTW)_{df}$ com a distância DTW. $CID(DTW)_{ori}$ obteve uma acurácia inferior que DTW em oito conjuntos de dados e similar em nove. Já $CID(DTW)_{df}$ obteve uma acurácia inferior em apenas dois conjuntos de dados, e similar em oito.

Em linhas gerais, podemos tirar as mesmas conclusões da comparação de CID com ED. Novamente foi possível verificar, de maneira independente e utilizando uma abordagem diferente que a invariância a complexidade é um fator importante a ser considerado na classificação de séries temporais, e que há espaço para melhora com relação a $CID(DTW)_{orig}$.

Na Figura 10 é apresentada uma comparação entre $CID(DTW)_{ori}$ e $CID(DTW)_{df}$. Assim como no caso em que ED era usada como distância base, essa figura mostra que $CID(DTW)_{df}$ é competitivo com $CID(DTW)_{ori}$. Em 22 conjuntos de dados, $CID(DTW)_{df}$ obteve uma acurácia melhor que CID_{ori}, e houve seis empates.

A utilização da dimensão fractal como fator de correção de complexidade também tem outra vantagem sobre a formulação original. A dimensão fractal é independente do

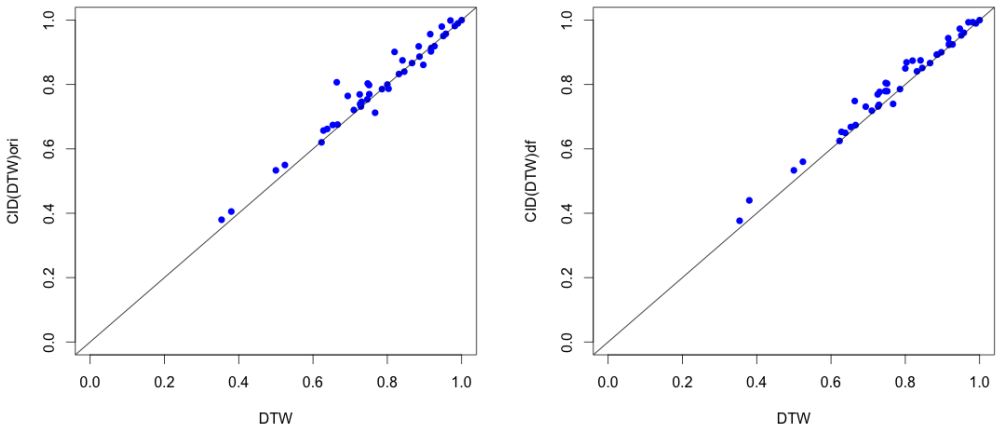

Figura 9. esquerda) comparação de desempenho entre $CID(DTW)_{ori}$ e a distância DTW; direita) comparação de desempenho entre $CID(DTW)_{df}$ e a distância DTW. Pontos acima da diagonal indicam que CID é melhor que DTW

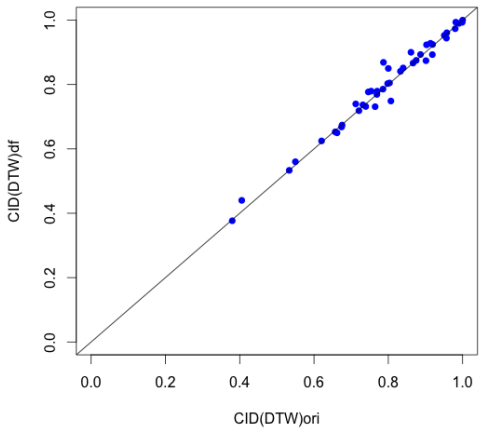

Figura 10. Comparação de desempenho entre $CID(DTW)_{ori}$ e $CID(DTW)_{df}$. Pontos acima da diagonal indicam que $CID(DTW)_df$ é melhor que $CID(DTW)_{ori}$

tamanho da série, enquanto a formulação original não é aplicavel para séries de tamanhos diferentes. Na Figura 11 é apresentado o diagrama de diferenças críticas com relação aos métodos baseados em dimensão fractal para calcular a CID, gerado a partir do teste de Friedman seguido do test post-hoc de Nemenyi, com nível de confiança de 95%. O método baseado em wavelets foi o que apresentou melhor resultado (em média), mas não foi estatisticamente melhor que o variograma, madograma e hallwood.

Os resultados experimentais apresentados até agora tem um viés positivo para os resultados obtidos a partir da dimensão fractal, pois utilizamos o melhor resultado obtido no conjunto de teste. Para contornar esse viés, realizamos um experimento com o intuito de tentar prever, usando apenas o conjunto de treinamento, qual seria o melhor método

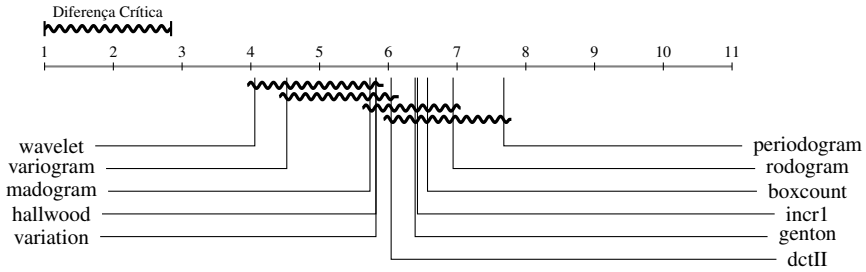

Figura 11. Diagrama de diferenças críticas obtido pelo teste de Friedman seguido pelo teste post-hoc Nemenyi, com nível de confiança de $0,95$. Valores à esquerda tem um desempenho (em média) melhor. Não há diferenças significativas entre métodos unidos por uma linha ondulada.

para estimar a complexidade da séria utilizando dimensão fractal. Esse experimento consiste em utilizar a técnica *leave-one-out* no conjunto treino para escolher qual seria o melhor método. Na Figura 12 estão sumarizados os resultados desse experimento. Dos 24 conjuntos de dados em que a melhor CID_{df} foi igual ou superior a CID_{ori}, o melhor método escolhido a partir do conjunto de treinamento foi superior ou igual em 17 (70% dos casos). Dos 28 conjuntos de dados em que a melhor $CID(DTW)_{df}$ foi igual ou superior a $CID(DTW)_{ori}$, o melhor método escolhido a partir do conjunto de treinamento superior ou igual em 21 (75% dos casos).

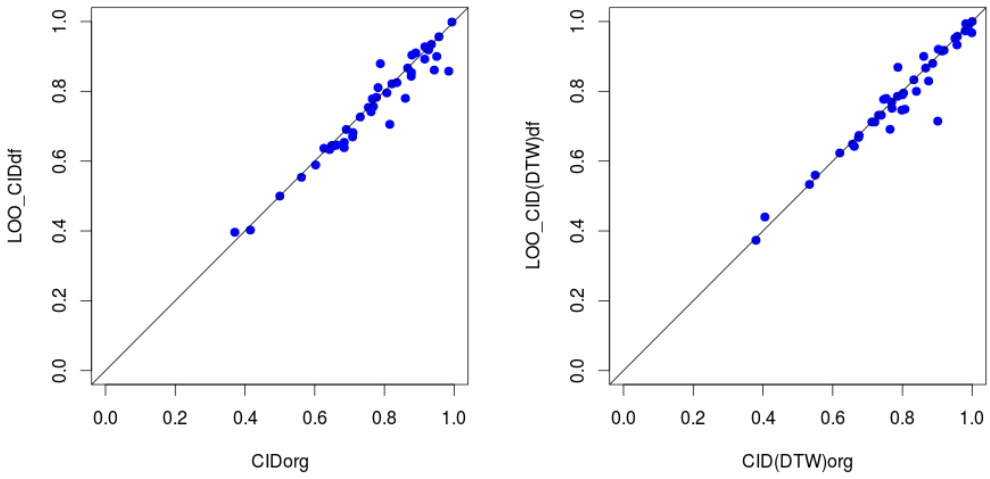

Figura 12. esquerda) comparação de desempenho entre CID_{ori} e CID_{df}, utilizando o melhor método predito no conjunto de treino; direita) comparação de desempenho entre $CID(DTW)_{ori}$ e $CID(DTW)_{df}$, utilizando o melhor método predito no conjunto de treino. Pontos acima da diagonal indicam que CID_{df} é melhor que CID_{ori}

5. Considerações Finais

Neste trabalho propusemos o uso de dimensão fractal como estimativa de complexidade de séries para derivar uma medida de distância invariante à complexidade. Resultados

experimentais mostram que o uso da dimensão fractal é competitivo com a medida de complexidade originalmente proposta em [Batista et al. 2011]. Além disso, comprovamos empiricamente, utilizando uma abordagem diferente da originalmente proposta, que invariância à complexidade é um aspecto relevante na classificação de séries temporais. A utilização da dimensão fractal é uma alternativa mais geral do que a originalmente proposta pois ela não requer que as séries tenham a mesma amplitude para serem comparadas.

Agradecimentos Trabalho desenvolvido com auxílio a pesquisa da FAPESP.

Referências

Batista, G. E. A. P. A., Wang, X., and Keogh, E. J. (2011). A complexity-invariant distance measure for time series. In *Proceedings of the Eleventh SIAM International Conference on Data Mining, SDM 2011*, pages 699–710. SIAM / Omnipress.

Chandola, V., Cheboli, D., and Kumar, V. (2009). Detecting anomalies in a time series database. Technical Report 09-004, Computer Science Dep., U. of Minnesota.

Ding, H., Trajcevski, G., Scheuermann, P., Wang, X., and Keogh, E. J. (2008). Querying and mining of time series data: experimental comparison of representations and distance measures. *Proceedings of the VLDB Endowment*, 1(2):1542–1552.

Faloutsos, C., Ranganathan, M., and Manolopoulos, Y. (1994). Fast subsequence matching in time-series databases. In *ACM SIGMOD International Conference on Management of Data*, pages 419–429. ACM Press.

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software*, 31(7):1–24.

Gneiting, T., Ševčíková, H., and Percival, D. B. (2010). Estimators of fractal dimension: Assessing the roughness of time series and spatial data. Technical Report 577, University of Washington, Department of Statistics.

Keogh, E., Zhu, Q., Hu, B., Hao, Y., Xi, X., Wei, L., and Ratanamahatana, C. A. (2011). The UCR time series classification/clustering homepage. `www.cs.ucr.edu/~eamonn/time_series_data/`.

Keogh, E. J. (2003). Efficiently finding arbitrarily scaled patterns in massive time series databases. In *European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'2003)*, volume 2838 of *LNCS*, pages 253–265. Springer.

Keogh, E. J., Wei, L., Xi, X., Vlachos, M., Lee, S.-H., and Protopapas, P. (2009). Supporting exact indexing of arbitrarily rotated shapes and periodic time series under euclidean and warping distance measures. *VLDB J.*, 18(3):611–630.

Li, K.-C., Yan, M., and Yuan, S. (2002). A simple statistical model for depicting the cdc15-synchronized yeast cell-cycle regulated gene expression data. *Statistica Sinica*, 12:141–158.

Zhu, Q., Rakthanmanon, G. B. T., and Keogh, E. (2012). A novel approximation to dynamic time warping allows anytime clustering of massive time series datasets. In *12th SIAM International Conference on Data Mining*.

Zunic, J. D., Rosin, P. L., and Kopanja, L. (2006). Shape orientability. In *Computer Vision - ACCV 2006*, volume 3852 of *LNCS*, pages 11–20. Springer.