

### Projeto 3 - Resultados

Nesse projeto, o principal objetivo era gerar três diferentes padrões de ondas por autômatos e tentar encontrar medidas que possibilitem a separação de cada um dos três conjuntos.

As medidas utilizadas nesse trabalho foram as seguintes:

- Desvio padrão das frequências relativas
- Entropia
- Paridade
- Distância entre símbolos
- Tamanho de *Burst*
- Frequência máxima do espectro de potência
- Número de conexões (grafo)
- Razão entre conexões realizadas e possíveis
- Área

Iniciamos o projeto a partir da implementação de uma função que recebe como entrada a matriz estocástica e a quantidade de passos e nos retorna um vetor com os dados do autômato gerado. As matrizes 1, 2 e 3 foram utilizadas em nossa função com 300 passos e geraram respectivamente as Figuras 1, 2 e 3

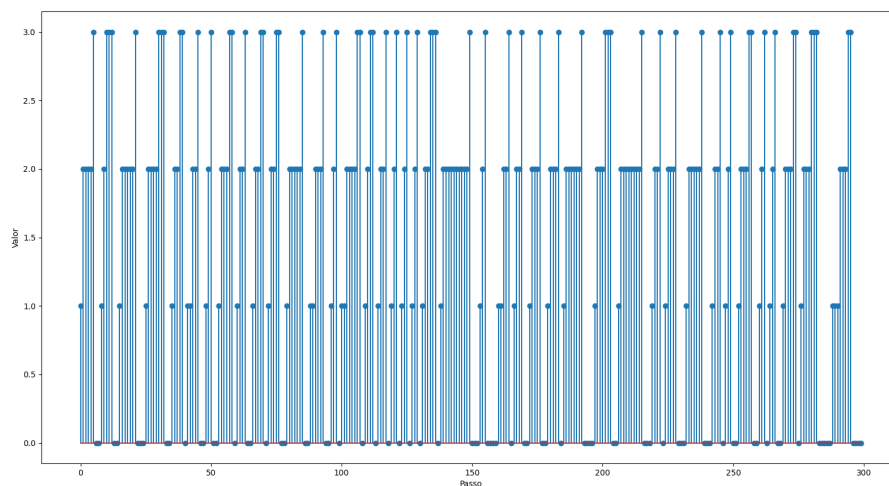


Figura 1: Onda dente de serra

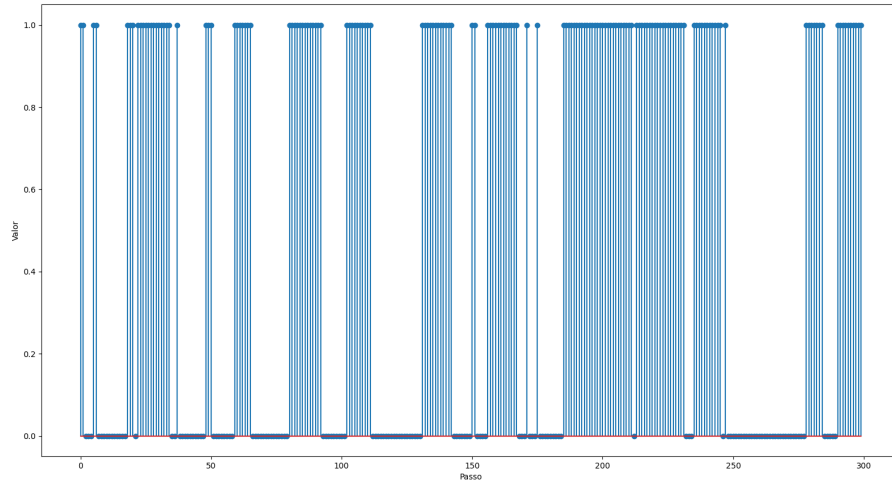


Figura 2: Onda retangular

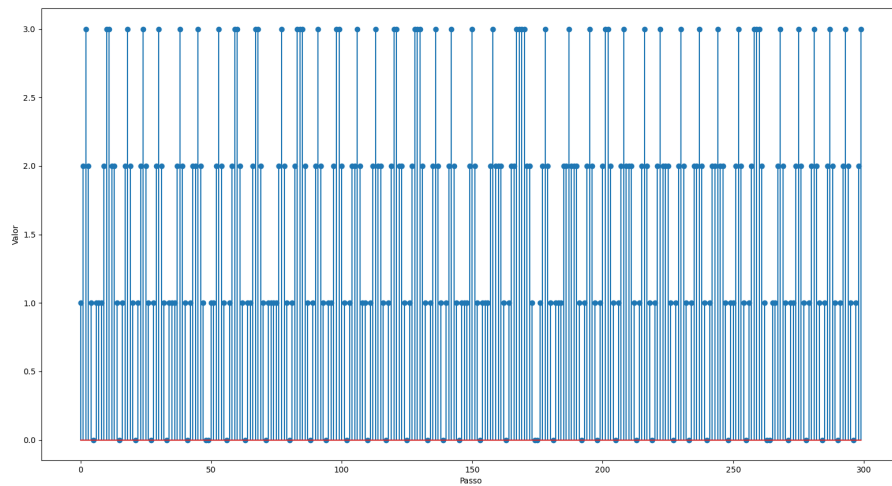


Figura 3: Onda triangular

|   | 0   | 1   | 2   | 3   |
|---|-----|-----|-----|-----|
| 0 | 0.5 | 0   | 0   | 0.7 |
| 1 | 0.5 | 0.1 | 0   | 0   |
| 2 | 0   | 0.9 | 0.6 | 0   |
| 3 | 0   | 0   | 0.4 | 0.3 |

Tabela 1: Matriz estocástica onda dente de serra

Com tal gerador, seguimos para a construção do banco de dados de 150 vetores com 50 autômatos de cada tipo. Com o intuito de separá-los, iniciamos as análises a partir do cálculo da

|   |     |     |
|---|-----|-----|
|   | 0   | 1   |
| 0 | 0.9 | 0.1 |
| 1 | 0.1 | 0.9 |

Tabela 2: Matriz estocástica onda retangular

|   |     |     |     |     |     |     |
|---|-----|-----|-----|-----|-----|-----|
|   | 0   | 1   | 2   | 3   | 4   | 5   |
| 0 | 0.1 | 0   | 0   | 0   | 0   | 0.9 |
| 1 | 0.9 | 0.2 | 0   | 0   | 0   | 0   |
| 2 | 0   | 0.8 | 0.1 | 0   | 0   | 0   |
| 3 | 0   | 0   | 0.9 | 0.3 | 0   | 0   |
| 4 | 0   | 0   | 0   | 0.7 | 0.2 | 0   |
| 5 | 0   | 0   | 0   | 0   | 0.8 | 0.1 |

Tabela 3: Matriz estocástica onda Triangular.

frequência relativa de cada um dos símbolos presentes no vetor analisado. Criamos uma função que calcula a frequência relativa de cada um dos símbolos do vetor e para retirarmos um único número que caracterize esse vetor, utilizamos o desvio padrão dessas frequências relativas.

As segunda e terceira análises estão relacionadas à entropia e a paridade do vetor seguindo o cálculo da entropia de Shannon em bits. Como cada uma dessas análises já é um único valor, utilizamo-as diretamente como atributo do vetor.

A quarta medida é o valor médio da distâncias entre símbolos no vetor. Para isso, implementamos uma rotina que encontra a média das distâncias entre todos os símbolos do vetor. Essa rotina então nos retorna alguns valores, um para cada símbolo. Nosso parâmetro baseado nas distâncias entre símbolos, foi definido como a média desses valores.

A quinta medida foi o tamanho dos *bursts* de cada símbolo do vetor. Novamente, temos uma série de valores para cada um dos símbolos, de forma que definimos o parâmetro para comparação como a média dos tamanhos de *bursts* de todos os símbolos.

A sexta aferição foi baseada no espectro de potência de nossos vetores. Identificamos então a frequência que mais se destacava em cada vetor como nosso parâmetro. A transformada de Fourier foi feita utilizando a implementação da biblioteca numpy. Um exemplo dessa medida pode ser encontrado na Figura 4

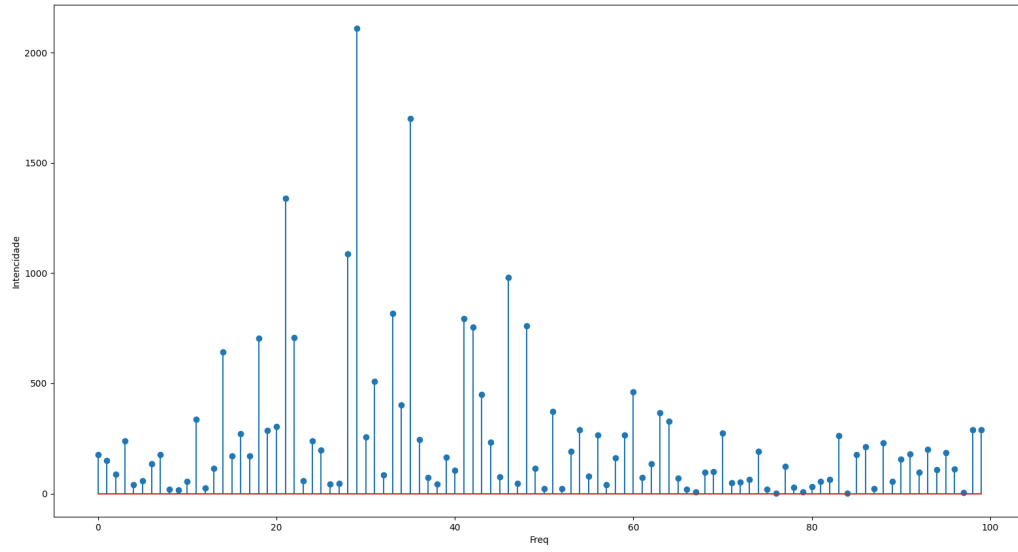


Figura 4: Representação do espectro de potência de um autômato de onda dente de serra.

Seguindo para a representação em grafos de nosso sistema, calculamos grau de cada nó assim como a razão entre o número de conexões ocorridas e o número possível de conexões. Um exemplo de matriz de adjacência obtida pode ser visto na Figura 5

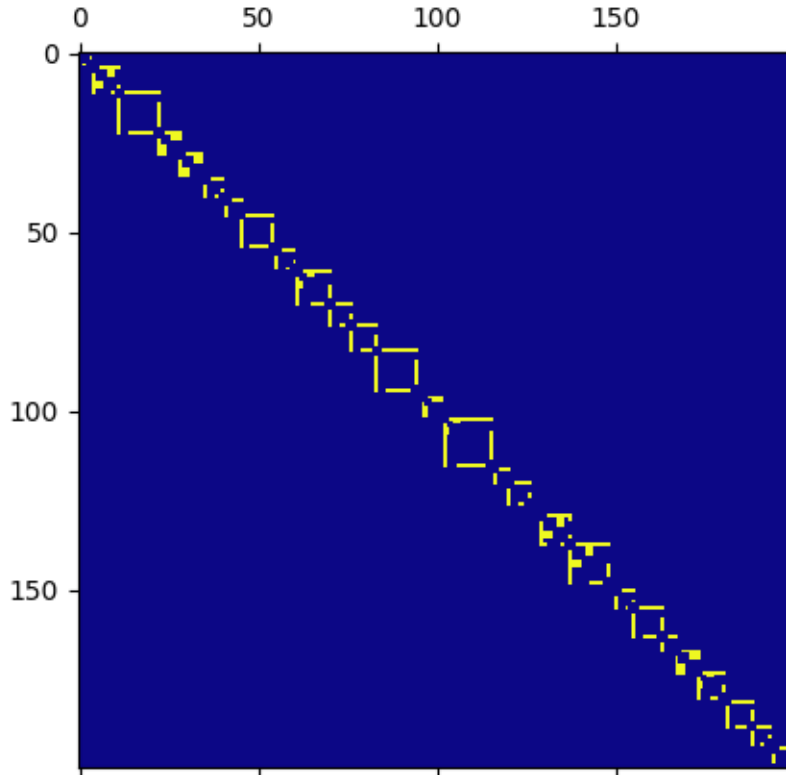


Figura 5: Matriz de adjacência obtida pela função implementada

Por fim, utilizamos o cálculo da área de cada curva a partir da integração dos vetores gerados como o ultimo atributo.

Para determinar se tais parâmetros obtiveram sucesso em separar os conjuntos abordados, fizemos diversos *scatterplots* relacionando-os, dois a dois. Tais gráficos podem ser observados na Figura 6. Nessa figura, fica claro que alguns parâmetros são similares devido à sua origem, por exemplo a entropia e a paridade, que é calculada como dois elevado à entropia do vetor em questão. Outro caso de similitude são os parâmetros dos grafos relacionados ao número de conexões e a razão de conexões realizadas e conexões possíveis.

Seguindo pela avaliação visual, chegamos à conclusão de que a melhor dupla de parâmetros que separa nossos conjuntos é o do número de conexões com a distância entre símbolos. Podemos notar que as demais duplas não encontram dificuldades em separar a onda retangular das outras duas, porém a separação entre as ondas dente de serra e triangular se mostras desafiadora dada a geometria semelhante entre tais ondas. Além disso, vemos a alta correlação entre alguns parâmetros (como já mencionado, entropia e paridade e o número de conexões e a razão de conexões realizadas e conexões possíveis) que apresentam formações aproximadas a retas.

Por fim, fica evidente a partir desses plots que a escolha de cada um dos parâmetros utilizados para a separação dos conjuntos é de extrema importância, pois alguns dos parâmetros escolhidos foram redundantes e resultam em gráficos extremamente semelhantes, adicionando

pouca ou nenhuma informação relevante para a separação dos dados. Esses parâmetros redundantes são geralmente modificações por um fator de escala em um dos parâmetros pré existentes.

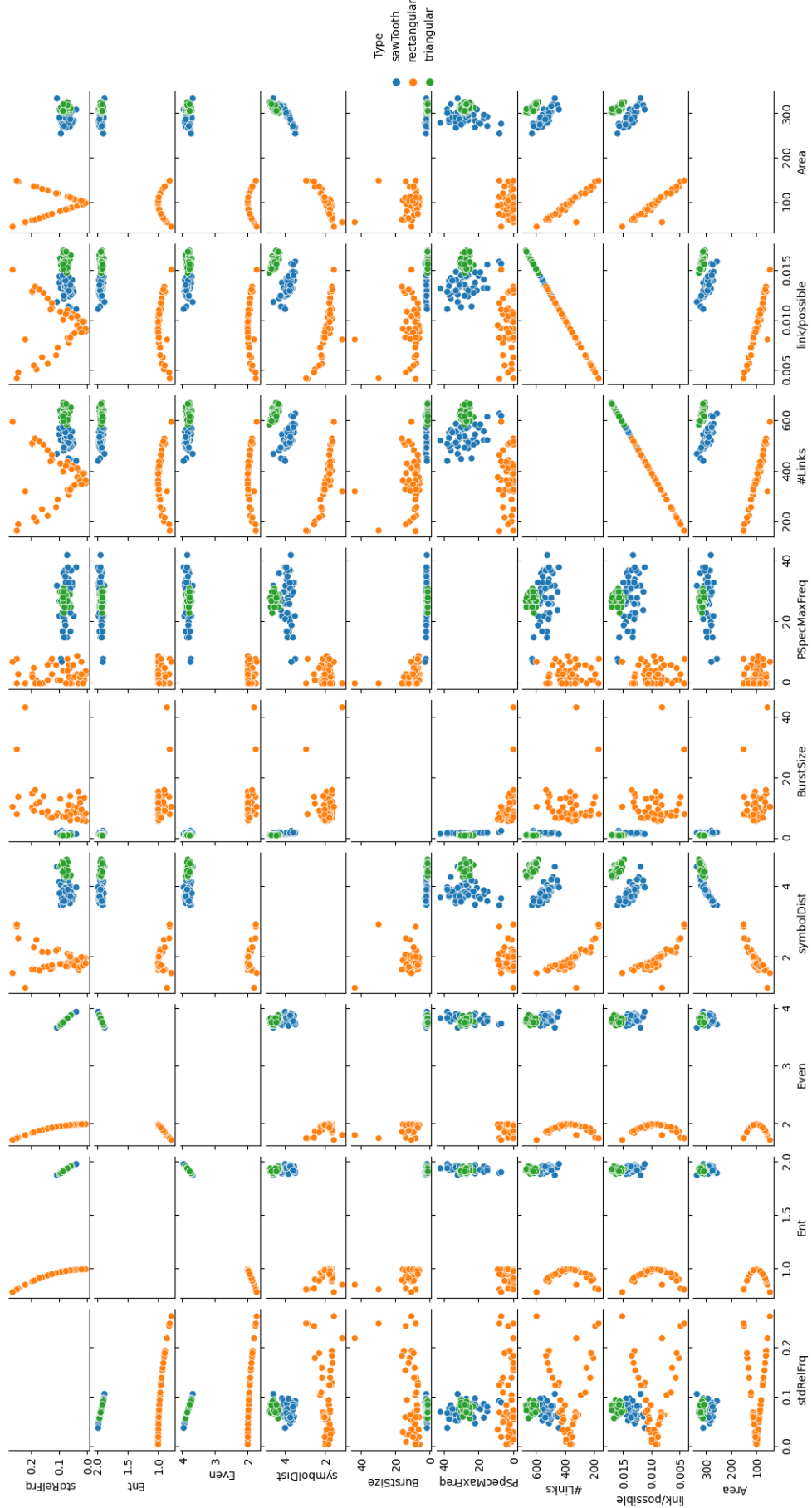


Figura 6: Gráfico de barras dos dados disponibilizados