



**UnB**

**CIC0203 - Computação Experimental -  
TA - 2022.2 - Tarefa T4 - Análise  
Bibliométrica**

URL Read-only Overleaf: <https://www.overleaf.com/read/rbvrrxytdqwm>

Paulo Mauricio Costa Lopes (RequiemDosVivos)

Brasília, 2023-01-29 23:04:06Z



# Lista de tarefas pendentes



# Sumário

<b>I</b>	<b>Preparação</b>	<b>3</b>
<b>II</b>	<b>Pesquisa Bibliométrica</b>	<b>5</b>
<b>1</b>	<b>T4 - Análise Bibliográfica sobre Simulação e Ensino-Aprendizagem, por Paulo Mauricio Costa Lopes</b>	<b>7</b>
1.1	Planejamento do Estudo . . . . .	7
1.1.1	O que já existe de pesquisa bibliométrica sobre esse tema? . . . . .	7
1.1.2	Uso do Bibliometrix e Biblioshiny . . . . .	8
1.1.3	Limitações . . . . .	8
1.2	Coleta de dados . . . . .	8
1.2.1	Query de Busca . . . . .	8
1.2.1.1	Explicação para os termos de busca usados . . . . .	8
1.2.2	Registros recuperados . . . . .	9
1.3	Análise dos dados . . . . .	9
1.3.1	Filtragem de registros . . . . .	9
1.3.2	Análise descritiva do dataset . . . . .	9
1.4	Visualização dos dados . . . . .	10
1.4.1	<i>Annual-Scientific Production</i> . . . . .	10
1.4.2	<i>Three-Field Plots</i> . . . . .	10
1.4.3	<i>Average citations per Year</i> . . . . .	12
1.4.4	<i>Palavras Relevantes</i> . . . . .	13
1.5	Interpretação dos dados . . . . .	14
1.5.1	Annual Scientific Production . . . . .	14
1.5.2	Three-Field Plots . . . . .	14
1.5.3	Average Citations per Year . . . . .	14
1.5.4	Palavras Relevantes . . . . .	14

## *SUMÁRIO*

# Lista de Figuras

1.1	Evolução da produção científica . . . . .	10
1.2	Three-Field Plot - I . . . . .	11
1.3	Three-Field Plot - II . . . . .	11
1.4	Three-Field Plot - III . . . . .	12
1.5	Media de Citações por Ano . . . . .	12
1.6	Palavras Relevantes . . . . .	13
1.7	Word Cloud . . . . .	13

## *LISTA DE FIGURAS*



# Lista de Tabelas

# Resumo

Este documento contém o produto da tarefa especificada no título deste documento, conforme as orientações em <https://www.overleaf.com/read/cytswcjsxxqh>.



# Parte I

## Preparação



# Parte II

## Pesquisa Bibliométrica



# Capítulo 1

## T4 - Análise Bibliográfica sobre Simulação e Ensino-Aprendizagem, por Paulo Mauricio Costa Lopes

### 1.1 Planejamento do Estudo

O modelo que será objeto de estudo neste trabalho é o de **virus\_on\_network** “A virus model with some number of agents entering in contact through a network of relations”, um modelo multi-agente de conexão e infecção para COVID-19. Esse trabalho foi adaptado para ser usado em infecção de conectividade de computadores.

Esse trabalho busca responder às seguintes perguntas:

- Como a simulação multiagente tem sido usada em sistema de infecções?
- Como podemos adaptar esse trabalho?
- Qual a importância desse modelo?
- Quais os principais temas e soluções ligadas a modelos similares?

#### 1.1.1 O que já existe de pesquisa bibliométrica sobre esse tema?

Com base nos resumos lidos, os artigos que analisam o problema "Infecção e propagação de malware" com inúmeras abordagens: usando simulação de multi-agent ou individual-based models; discorrendo sobre a caracterização dos indivíduos da rede, IoT, mobile, computadores pessoais e redes heterogêneas; A topologia da rede também é algo levado em consideração, p2p, bus, mesh; Por fim alguns artigos tentam analisar os métodos de se defesa.



### 1.1.2 Uso do Bibliometrix e Biblioshiny

Serão usadas a ferramenta e o *workflow* proposto pelos autores do pacote Bibliometrix, conforme indica a figura ??.

### 1.1.3 Limitações

Inicialmente meu tema escolhido tinha sido Schelling - modelo de segregação, já havia feito algumas atividades e após algumas tarefas tivemos um incidente que alguém apagou o repositório/algumas tarefas no overleaf da materia tornando mais trabalhoso enviar as tarefas. Algumas aulas apos o ocorrido levantei a ideia de trocar meu tema para , ele falou que seria necessário algumas adaptações mais que era possível. Então meu tema passou a ser **virus\_on\_network** e estou refazendo as atividades anteriores.

## 1.2 Coleta de dados

A coleta de dados feita utilizando o Web of Science, as consultas foram feitas inicialmente nos dias 15 de Novembro de 2022 mas devido a troca de tema realizei novas buscas nos dias 8 de Janeiro de 2023 e 15 de Janeiro de 2023, acessado pelo Portal de Periódicos da CAPES. Com a nova query foram retornados 763 artigos.

### 1.2.1 Query de Busca

A query usada inicialmente foi:

ALL=( ( (agent-based) OR (simulation model))AND ( (propagation) OR (proliferate)) AND ( (malware) OR (virus) OR (worm) ))

Porém essa query retornava artigos relacionados a saúde então foi feita uma nova query para que os artigos resultantes fossem somente relacionados área de tecnologia.

ALL=( ( ( agent-based ) OR ( model ) OR ( simulation ) ) AND ( propagation ) AND ( malware ) OR ( worm ) ) AND (network) )

[Link para query](#)

#### 1.2.1.1 Explicação para os termos de busca usados

A busca consistiu de quatro conjuntos de termos ligados por *AND*, aplicadas à busca por tópico (O termo de busca pode aparecer no Título, no Resumo, na Author Keywords, ou nas Keywords Plus da referência)

Primeiro conjunto de termos: Os termos **agent-based**, **model** e **simulation** foram usados para recuperar artigos que tenham em seu título, palavras-chave e resumo, termos relacionados a métodos modelos ou simulações.

Segundo conjunto de termos: O conjunto unitário **propagation**, inicialmente tinha sido um conjunto com dois elementos mas após teste resolvo torna-lo um conjunto unitário para aprimorar a busca.

Terceiro conjunto de termos: Os termos **malware** e **worm** foram utilizados para selecionar artigos que fossem relacionados a malwares a worms (um tipo de malware específico).

Quarto conjunto de termos: O quarto conjunto de termos é um conjunto unitário *Network* que foi utilizado para explicitar que essas infecções ocorreram por network.

### 1.2.2 Registros recuperados

Os 763 registros obtidos como resultado da busca encontram-se em <https://github.com/jhcf/Comput-Experim-20212-Overleaf2/exploratory-data-analysis/RequiemDosVivos/PesqBibliogr/WoS2022/resic.txt>, Foram utilizadas as opções *Exportar registros para arquivo de texto sem formatação e registro completo e referencias citadas* no WoS, para que as citações também fosse usadas em análises da citações (estrutura intelectual do conhecimento). Os 763 registros foram recuperados em dois blocos de até 500 registros por vez (1-500, 501-763).

## 1.3 Análise dos dados

### 1.3.1 Filtragem de registros

Foi aplicado um filtro ao dataset inicial, com 763 registros, que continham prévias de artigos, artigos de conferência, capítulos de livro etc. Foram mantidos apenas os registros de artigos publicados em revistas científicas e foram excluídos os que estavam em chins (sobraram somente os artigos em inglês). Após a aplicação desse filtro, 744 registros foram mantidos no dataset.

### 1.3.2 Análise descritiva do dataset

As informações mais gerais sobre o dataset são as seguintes:

**Timespan** Os artigos que atenderam aos critérios de busca e filtragem foram publicados a partir de 1996 até 2023.

**Sources** São 457 fontes de informação que publicaram os documentos recuperados no dataset.

**Documents** Foram recuperados, como dito anteriormente, 744 documentos.

**Average citations per documents** Cada artigo no data foi citado, em média 45.51 vezes.

**References** O dataset contém 12907 referências citadas.

**Author's Keywords (DE)** 1539 distintas palavras-chave indicadas pelos autores foram encontradas no dataset.

**Authors** 1656 distintos nomes de autores foram encontrados no dataset

**Authors of single-authored documents** Dentre os 1656 distintos (nomes de) autores encontrados, 28 deles editaram artigos individualmente, isso é, sem co-autores.

**Co-Authors per Documents** A media de coautores por documentos foi de 3.49.

## 1.4 Visualização dos dados

### 1.4.1 *Annual-Scientific Production*

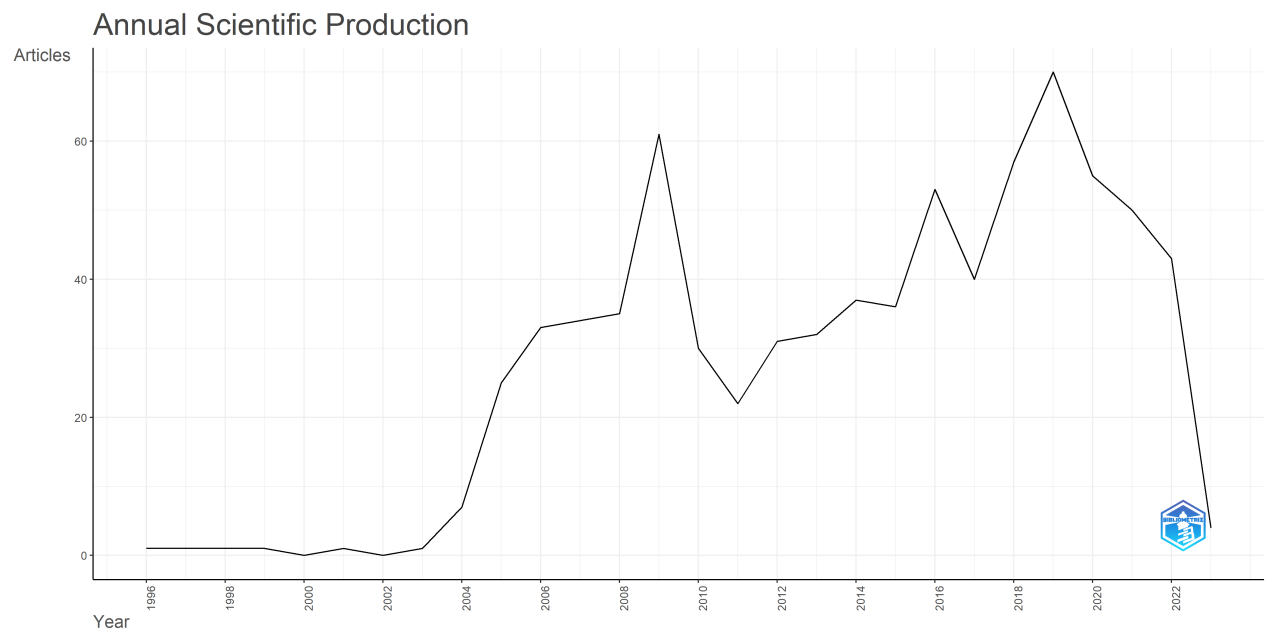


Figura 1.1: Evolução da produção científica

A figura 1.1 mostra a produção científica anual de 1996 até 2022 (e um pequeno espaço de tempo de 2023).

### 1.4.2 *Three-Field Plots*

As três figuras, 1.2, 1.3 e 1.4 mostram a relação país, revista e palavras chave, foram feitos 3 esquemas de imagem para melhor visualização das palavras-chave, revistas e países mais relevantes para o assunto.

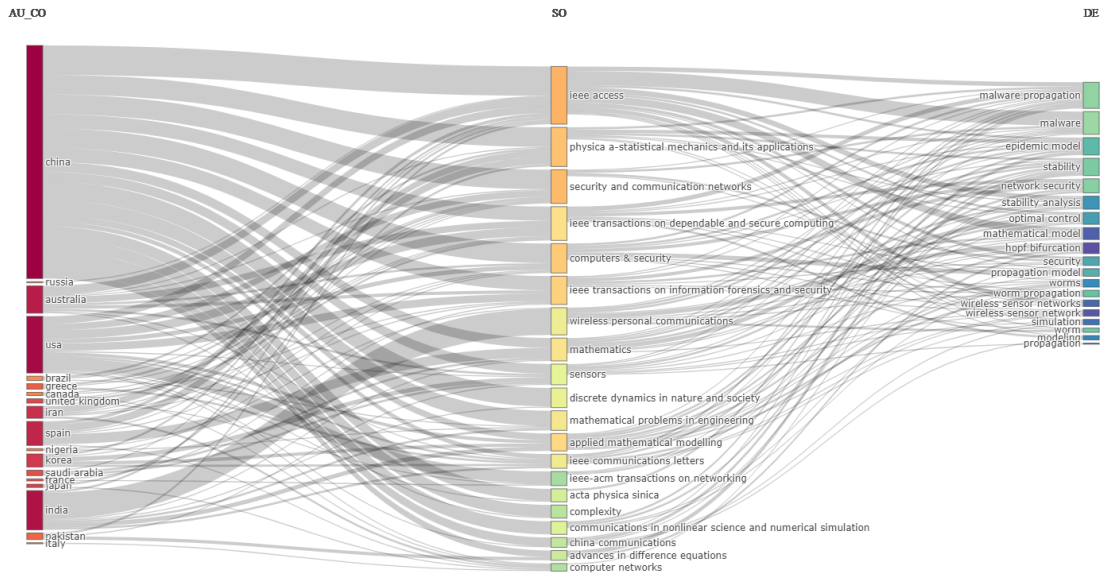


Figura 1.2: Three-Field Plot - I

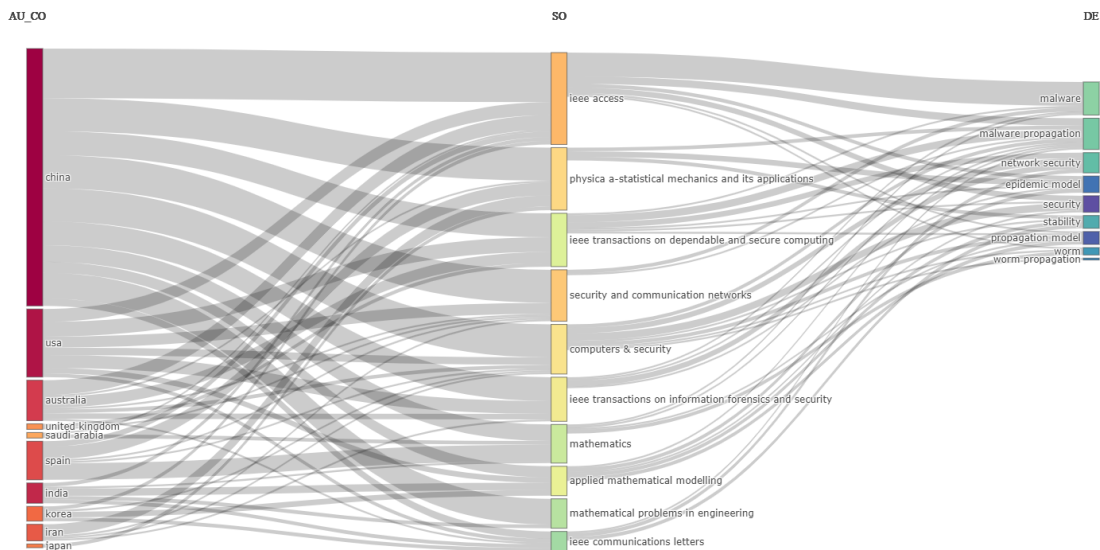


Figura 1.3: Three-Field Plot - II

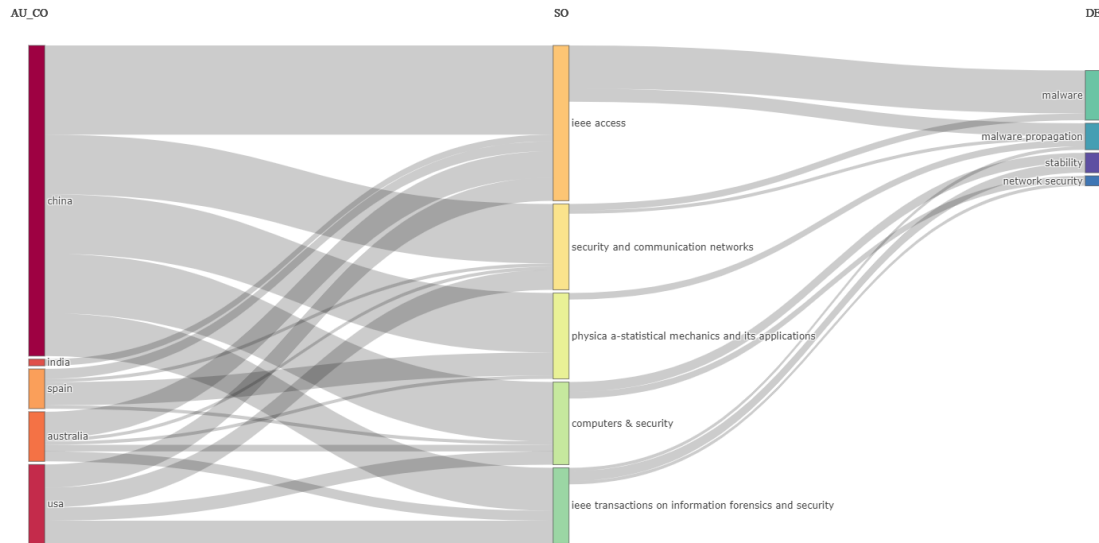


Figura 1.4: Three-Field Plot - III

### 1.4.3 *Average citations per Year*

A figura 1.5 representa graficamente a média de citações ao longo dos anos, de 1996 até 2023.

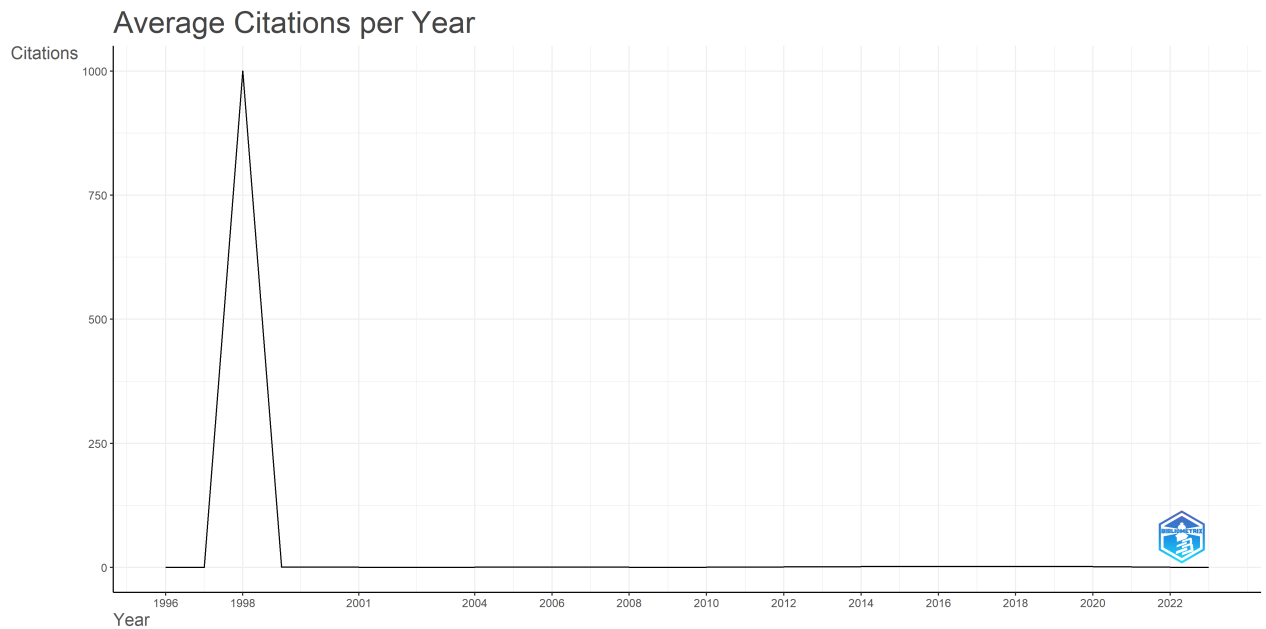


Figura 1.5: Media de Citações por Ano



## 1.5 Interpretação dos dados

### 1.5.1 Annual Scientific Production

Existem dois anos marcantes, o primeiro seria 2008 onde provavelmente teve o pico de produções científicas devido a popularização dos smartphones, lançamento dos primeiros celulares com Android. Após o ano de 2008 existe uma queda acentuada, que provavelmente foi devido a mudança da categoria dos smartphones (a priori **IoT** e agora **mobile**). O segundo ponto marcante pe 2019, onde a produção científica cresceu muito provavelmente devido o COVID-19, onde foi criado um ambiente mais propício para o uso e infecção de IoT.

### 1.5.2 Three-Field Plots

Inicialmente tinham sido feito plots com múltiplas imagens porém estava difícil visualizar e interpretar os dados então reduzi para 5 elementos, assim podemos inferir quais países produziram mais conteúdo, as revistas que mais publicaram e quais as palavras chaves mais relacionadas.

### 1.5.3 Average Citations per Year

Observando inicialmente a plotagem pensei que existia um erro nos dados então fui ler o artigo em questão, ele é realmente um artigo referenciado mais de 28 mil vezes. O artigo Collective dynamics of 'small-world' networks, de D J Watts, fala sobre o conceito de 'small-worlds' e como a organização é feita. Acredito que esse seja o motivo desse artigo ser tanto citado, muitos estudos devem basear as suas simulações nos conceitos de 'small-worlds' de Watts.

### 1.5.4 Palavras Relevantes

As palavras relevantes foram interessantes para visualização dos interesses dos autores. Apesar dos estudos serem diretamente relacionados a malware e worm, a preocupação dos autores parece estar mais voltada para **Propagação** e **transmissão**.