# Sistema de Privacidade Projeto 2 Relatório – 75150 e 72928

## Introdução

O objetivo deste projeto é analisar e melhorar a seleção de caminhos em circuitos Tor, focando na privacidade a nível geográfico.

O problema identificado é que circuitos inteiros podem estar localizados no mesmo país, aumentando o risco de deanonymização por autoridades nacionais. O algoritmo original do Tor apenas evita múltiplos nós no mesmo /16 subnet, o que não garante proteção contra adversários a nível de país.

## Implementação

### Parser

Foi implementado um parser (*ConsensusParser*) para ler documentos de consenso Tor e extrair dados de cada nó, incluindo:

Nickname, fingerprint, IP, portas (OR e DIR), flags, versão, largura de banda, política de saída e país (via GeoIP).

O parser armazena os nós como objetos *Node* para posterior utilização na seleção de circuitos.

### Algoritmo Base

O algoritmo original seleciona três nós por circuito: Guard, Middle e Exit.

O peso de cada nó é definido com base na largura de banda, e filtros garantem a ausência de repetição de famílias e subnets. A seleção é aleatória, ponderada pelos pesos.

### Algoritmo Geo-Aware

O algoritmo geográfico ajusta os pesos dos nós com base na diversidade de países:

- α: aplicado na seleção do Guard. Se o guard não partilha país com o Exit, o peso é multiplicado por 1 + α.
- β: aplicado na seleção do Middle. O peso depende do número de nós no circuito que partilham país com o candidato (0–3 multiplicadores).

Este ajuste mantém a ponderação por largura de banda, mas favorece a escolha de nós de países diferentes, evitando filtros rígidos.

## Resultados Experimentais

### Node Diversity

| Posição | Original | Geo-Aware |
|---|---|---|
| Guard | 840 | 856 |
| Middle | 872 | 891 |
| Exit | 702 | 717 |
| Global | 2048 | 2061 |

- O algoritmo geográfico apresenta maior diversidade em todas as posições e globalmente.

Entropy

| Posição | Original | Geo-Aware |
|---|---|---|
| Guard | 9,6158 | 9,6565 |
| Middle | 9,6917 | 9,7377 |
| Exit | 9,2717 | 9,3103 |
| Global | 10,7982 | 10,8216 |

- A entropia global e por posição é ligeiramente maior no algoritmo geográfico, indicando maior dispersão de nós por país.

Bandwidth stats

| Posição | Original | Geo-Aware |
|---|---|---|
| Min | 440 | 140 |
| Max | 130 000 | 98 000 |
| Média | 21 842,3 | 21 116,8 |

- O Geo-Aware seleciona circuitos com menor largura de banda mínima (140 vs 440), refletindo que alguns nós mais lentos são escolhidos para manter diversidade geográfica.
- O máximo também é inferior (98 000 vs 130 000), mostrando que o algoritmo original consegue ocasionalmente formar circuitos de maior capacidade.
- A média de largura de banda mantém-se semelhante entre os dois algoritmos, indicando desempenho comparável no geral.

Distribuição de Bandwidth por Circuito

O algoritmo Geo-Aware apresenta um valor mínimo de largura de banda inferior ao Original (140 vs 440), indicando que alguns circuitos incluem nós mais lentos devido à priorização de diversidade geográfica. O valor máximo também é menor (98 000 vs 130 000), enquanto a média se mantém próxima

(21 116,8 vs 21 842,3), mostrando que o desempenho médio global é semelhante.

No geral, o Geo-Aware gera circuitos com largura de banda mais consistente, enquanto o Original apresenta maior variabilidade (Fig.1).

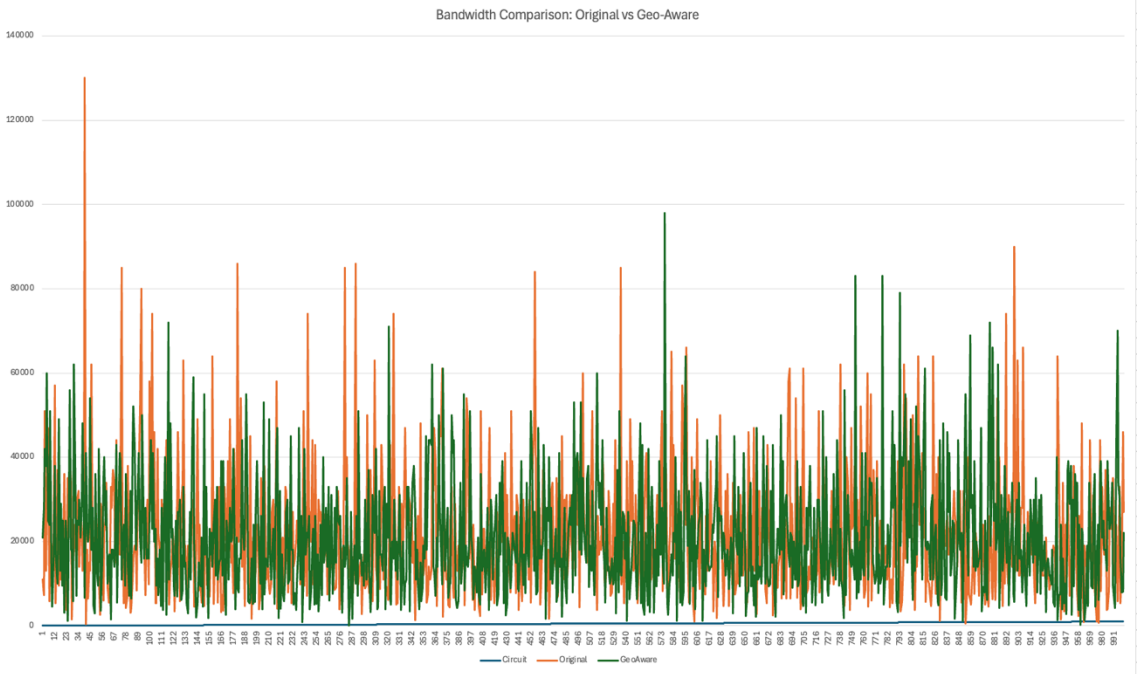


Fig1

**Conclusão**

A aplicação do algoritmo Geo-Aware permite um equilíbrio entre privacidade e desempenho, ao reduzir a concentração de nós do mesmo país em um circuito e melhorar a distribuição global de seleções.

Embora a prioridade seja a dispersão geográfica, o desempenho médio permanece comparável ao algoritmo Original, garantindo que a proteção adicional não compromete significativamente a largura de banda.

No conjunto, o Geo-Aware demonstra ser uma abordagem viável para aumentar a segurança dos circuitos Tor, oferecendo maior consistência e mitigando riscos de ataques de deanonymização sem necessidade de filtros rígidos.