

Prevendo Evolução de Exames BitTorrent com Redes Neurais Artificiais

Henrique Fan da Silva, discente de graduação,
Universidade Federal do Pampa, Campus Alegrete

Rodrigo Brandão Mansilha, docente,
Universidade Federal do Pampa, Campus Alegrete

henriquefan.aluno@unipampa.edu.br

A realização de monitoramento do universo de redes BitTorrent é desafiador pois exige um sistema distribuído com uma capacidade que deve ser proporcional ao sistema alvo (MANSILHA et al., 2011). Por exemplo, quanto maior o número de pares a serem monitorados, mais recursos são necessários. Além disso, devem ser previstos recursos adicionais para provimento de tolerância a falhas, cuja probabilidade de ocorrência aumenta conforme a duração do monitoramento (e.g. escala de semanas ou meses). Como o universo BitTorrent é dinâmico e suscetível a fenômenos de churn (i.e., entrada ou saída de pares em um taxa alta e um período curto de tempo), prever a evolução da quantidade de pares torna-se um problema chave para provisionamento adequado de recursos de monitoramento. Neste trabalho investiga-se o uso de Aprendizado Profundo para prever a evolução da quantidade de pares no exame ao longo do tempo, a partir dos dados do próprio monitoramento obtidos anteriormente. Para isso, modelamos a rede de monitoramento, e dividimos o tempo total de monitoramento em intervalos de tempo regulares (chamado de *snapshots*). Para realizar a predição, o algoritmo monta as matrizes de adjacência e características e as usam como entradas para uma Rede Neural Artificial (RNA). Por fim, explora-se o emprego de uma arquitetura de RNAs do estado da arte proposto. Para avaliar a proposta, seguimos a metodologia usada no algoritmo seminal (ZHAO et al., 2019). Essencialmente usamos um *dataset* e o separamos em uma parte de treinamento e outra parte de teste. Como referência usamos o conjunto de dados S1 para o qual possuímos bom grau de certeza sobre o tamanho do exame (CORDEIRO et al., 2021). Desse dataset (7200 *snapshots*) utilizamos 80% (5760) como conjunto de treinamento, e os outros 20% (1440) utilizados como conjunto de teste. Como métrica de desempenho comparamos a diferença entre o valor previsto e o valor medido, comparando o erro da predição com o modelo de média móvel integrado autorregressivo ARIMA, que é um dos modelos de série temporal mais populares e amplamente utilizados em diversas aplicações. Resultados preliminares indicam que a solução tem bom desempenho. Trabalhos em andamento e futuros incluem estudo de outros *datasets*, e comparação com técnicas mais simples de previsão.

Agradecimentos: Este trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), por meio do edital ARD 10/2020. Agradeço à FAPERGS por promover a inovação tecnológica, estimulando a

formação de recursos humanos e fortalecendo a expansão da infraestrutura da pesquisa no Estado.

Palavras-chave: Sistemas Distribuídos; Monitoramento; Redes Neurais Artificiais; Previsão.

Referências:

CORDEIRO, W. et al. Revisiting the coupon collector's problem to unveil users' online sessions in networked systems. *Peer-to-Peer Networking and Applications* , Springer, v. 14, n. 2, p. 687–707, 2021.

MANSILHA, R. B.; BAYS, L. R.; LEHMANN, M.; MEZZOMO, A.; FACCHINI, G.; GASPARY, L. P.; BARCELLOS, M. P.. Observing BitTorrent Universe Through Telescopes. In: 12th IFIP/IEEE International Symposium on Integrated Network Management (IFIP/IEEE IM 2011), Dublin, Ireland.

ZHAO, L. et al. T-gcn: A temporal graph convolutional network for traffic prediction. *IEEE Transactions on Intelligent Transportation Systems* , IEEE, v. 21, n. 9, p. 3848–3858, 2019