

Reinforcement Learning based Resource Management for Fog Computing Environment

Literature Review, Challenges, and Open Issues

Alunos:

Gabriel Freitas Willig

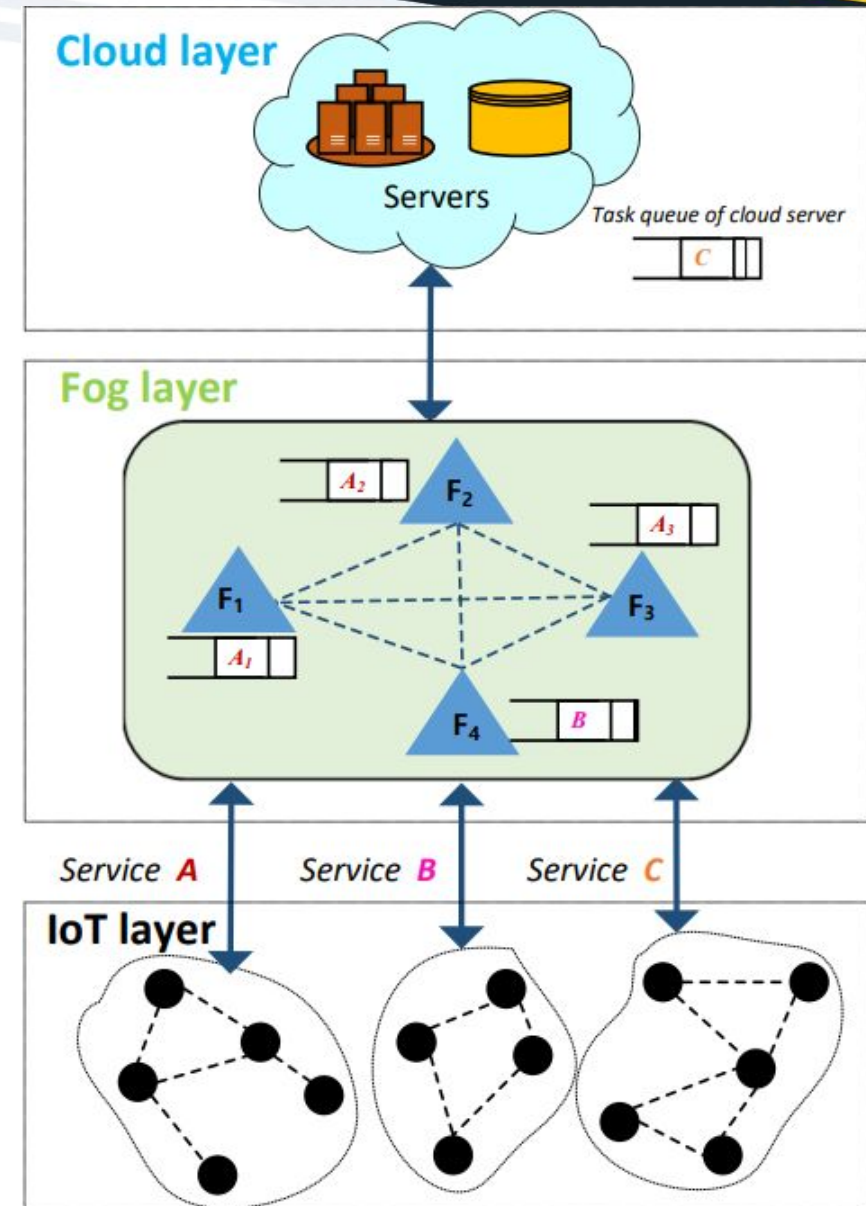
Luis Henrique Teodoro Lopes

Fog Computing

- Crescente da IoT.
- Trazer recursos para borda a fim de reduzir latência
- Elementos: Gateways, hubs, roteadores, VM's, switches, drones, carros e etc...
- Ambiente complexo, dinâmico e com recursos heterogêneos

Tipos

- **Centralizada**
 - Possui um controller e um conjunto de nodes
- **Hierárquica**
 - Master-slave
 - Federation (Resource sharing)
- **Distribuída**
 - Alocação de recursos é feita de forma distribuída
 - Fogs possuem uma tabela de recursos de cada fog vizinha



Fog Computing

Tabela de recursos

Mantém informações de CPU, Memória, RTT (Round-trip time), W (tempo de espera da fila) e f (frequência de clock)

Tarefa

$k = (A_k, B_k)$: A (Vetor de características da entrada), B (Vetor de recursos necessários)

Desafios

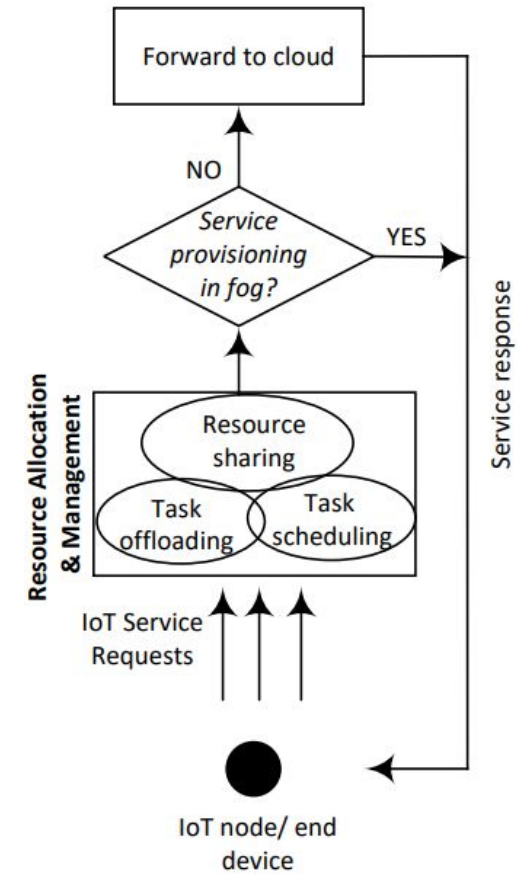
- Compartilhamento de recurso
- Escalonamento de tarefas
 - Qual recurso e em qual período de tempo irá processar a task
- Atribuição de tarefas
 - Qual nó irá processar

Solução Atual

- Sem framework bem definido
- Dependência de informação antecipada das tarefas
- Algoritmos com heurística
- Em sua maior parte, atacam somente um objetivo (ex: Consumo de Energia, Delay, e utilização dos recursos)

RESOURCE STATE TABLE OF NEIGHBORS OF FOG NODE F_1 .

Node ID	Fog specification & Resource Status				
	M	f	Processor	RTT	W
F_2	200	10	CPU	2.5	350.2
F_3	100	5	CPU & GPU	3.1	500
F_4	400	2.5	GPU	4.8	239.1

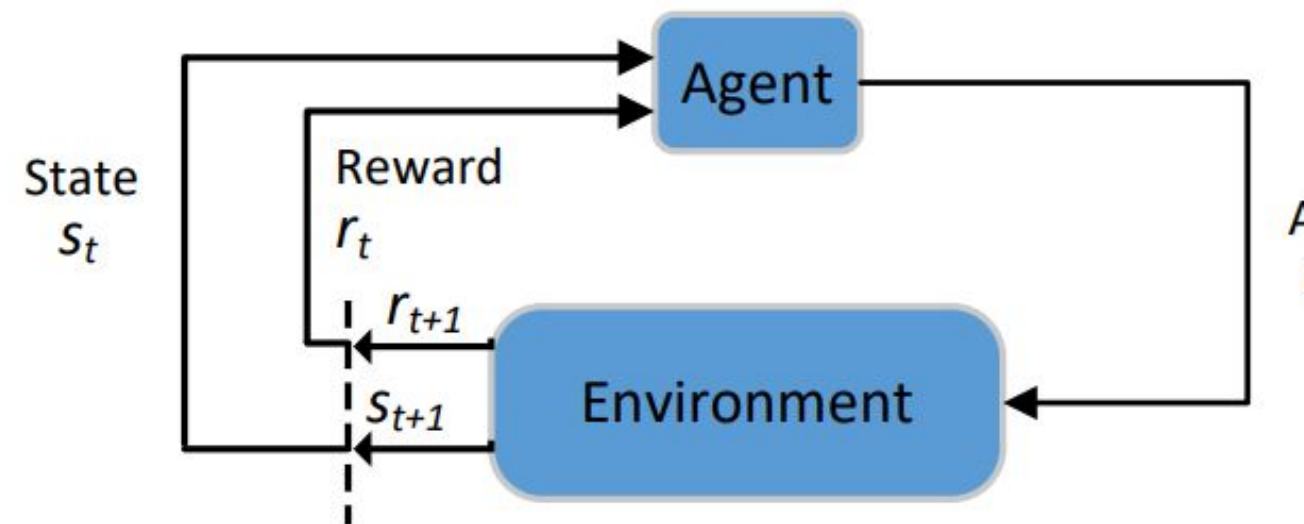


Reinforcement Learning

- Não é supervisionado nem não supervisionado
- Agentes aprendem um comportamento para tomada de decisão
- “Ciência da tomada de decisão”
- Se P e R são conhecidos
 - Princípio da Otimalidade de Bellman
 - Programação dinâmica. Resolução analítica de cada subproblema de otimização
- Raramente sabe-se a configuração do ambiente. Processo estocástico
 - MDP (Markovian Decision Process)
- Instância de RL: (S, A, R, P, γ)

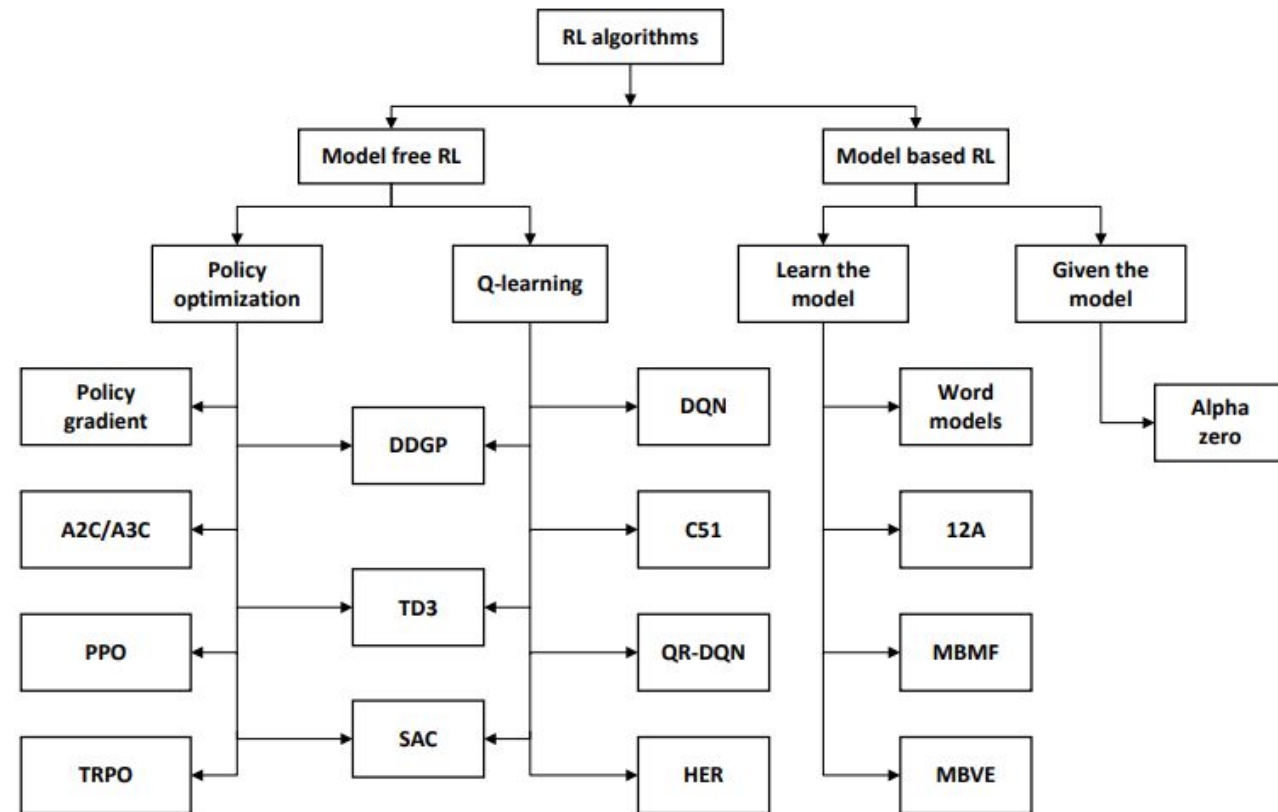
TABLE II
SYMBOLS USED IN THE WORK.

Symbol	Description
t	Time step t
a_t	Action that agent can take at t
s_t	State that agent is be in at t
r_t	Immediate reward for agent to take a_t
A	Action space, $A = \{a_t\}$
S	State space, $S = \{s_t\}$
R	Return or cumulative reward
Q	Q value (function)
π	Policy
γ	Discount factor
P	Transition probability



Reinforcement Learning

- Model-Free
 - Sem modelo prévio do ambiente
 - Simples e barato
 - Risco
 - Robôs
- Model-Based
 - Caro e Complexo
 - Modelo do ambiente
 - Carros autônomos



RL na Fog Computing

Três problemas principais: Resource Sharing, Task Scheduling e Task Offloading

Resource Sharing

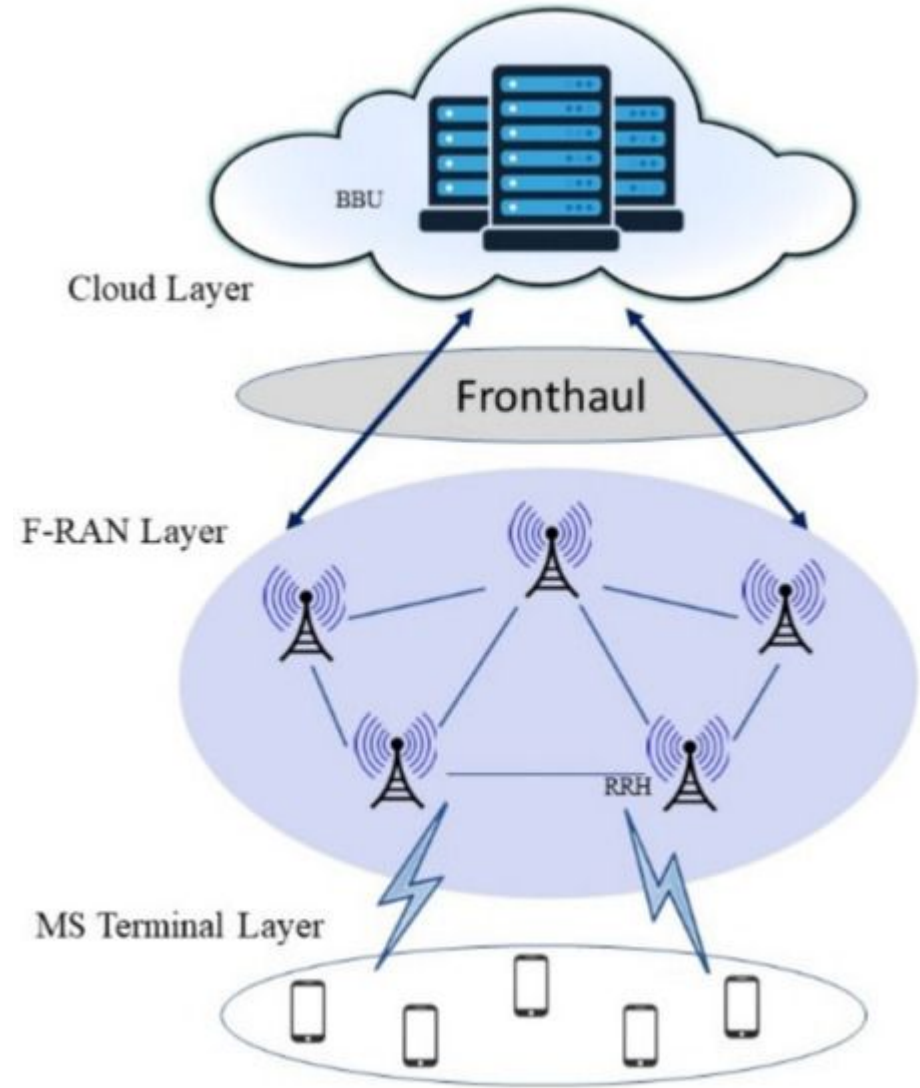
- Pool de recursos à serem consumidos

Caso F-RAN (Fog Radio Access Network)

- Modo de comunicação de um UE (Fog, Cloud ou D2D)
- On/Off states of Fog resources

Solução

- Recompensa: Negativo do consumo energético
- DRL (Deep Learning + Reinforcement Learning): Permite inputs não estruturados
- DRL dupla para tratar valores otimísticos de Q-value



RL na Fog Computing

Task Scheduling

Cenário IoT-fog-Cloud

Desafio de escolher onde uma task irá rodar: Camada da Fog, Cloud ou IoT

Schedulers (agentes) em GW (gateways) para alocação de recursos (VM)

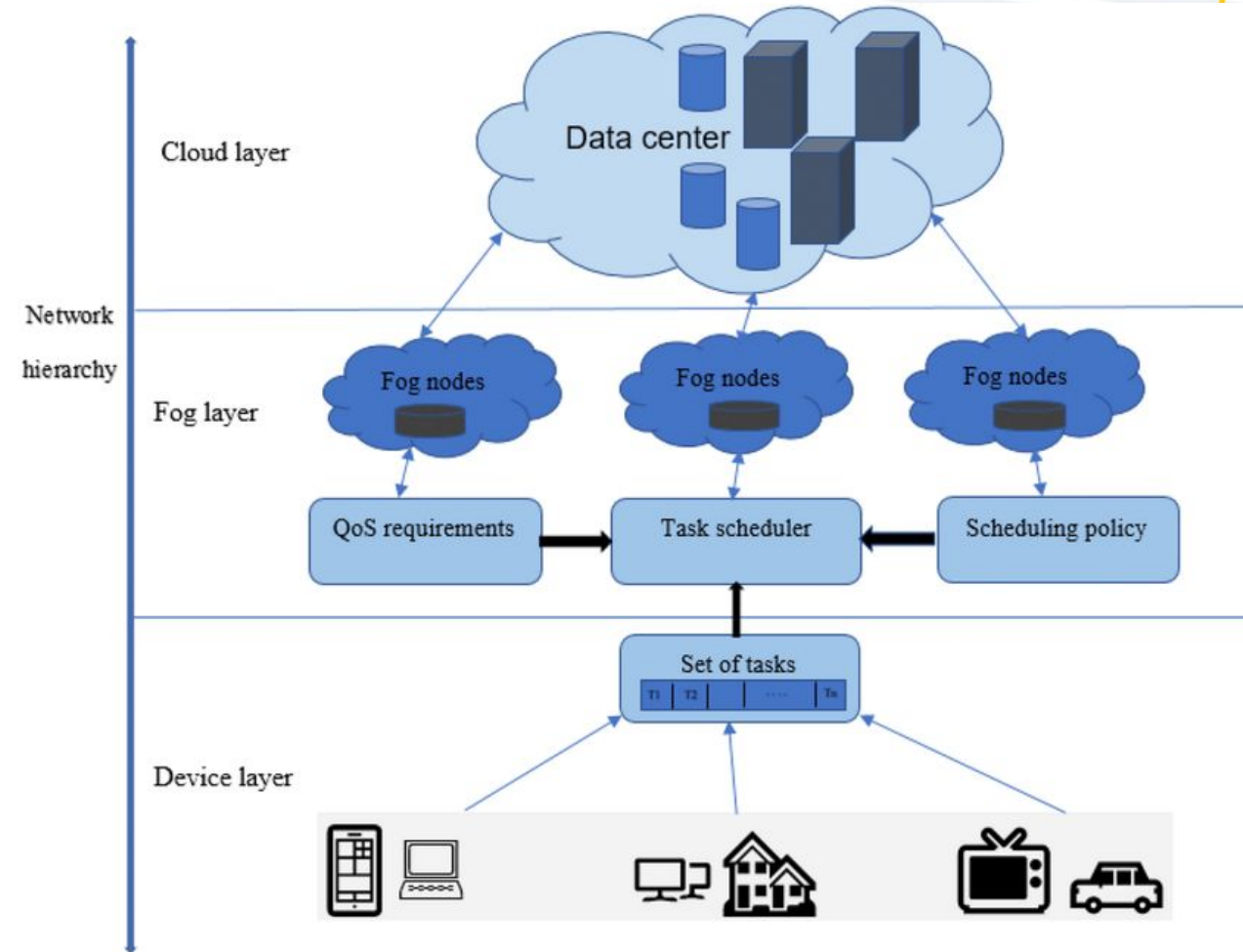
Entrada: $S = \{s(t) = (uCPU, uMemory, uStorage)\}$

$A = \{a \mid 1 \leq a \leq VM\}$ Em que é VM número total de VM's no caminho

Recompensa: Valor inverso ao delay de início da task

Solução

- DDQL (Double DQL)
 - Dois modelos Q1 e Q2 para reduzir a probabilidade de ações ineficientes
 - Acelera processo de aprendizado



Desafios e Problemas em Aberto

- **Relacionados à RL:**

- Nature of RL-based Algorithms: Necessidade de algoritmos eficientes e apropriados.
- Load Balancing: Dependência de políticas de exploração gera desequilíbrio.
- Task Scheduling: Heterogeneidade da capacidade dos dispositivos na camada de fog.
- Energy Efficiency Trade-off: Treinamento de modelos requer alto consumo de energia
- Real-time Responsiveness: Necessita tempo para processar recursos em larga escala.
- Security and Privacy: Capacidade limitada, necessidade de explorar utilização de blockchain
- Advance of Optimization Algorithms: Necessidade de reduzir a dimensão dos problemas para acelerar o processo de aprendizado.

Desafios e Problemas em Aberto

- **Relacionados a Fog Computing:**

- Resource Allocation in F-RAN: O RL precisa otimizar alocação de recursos heterogêneos sob demandas de baixa latência.
- Power Consumption Optimization for F-RAN: DRL enfrenta dificuldades para equilibrar economia de energia com requisitos em tempo real do F-RAN
- RL for Ultra-dense Fog Network: Falta de soluções de RL para particionar redes de fog em grande escala de forma eficiente
- Reliability of Fog Networks: Nós dinâmicos e falhas de máquinas virtuais comprometem QoS - RL ignora confiabilidade
- Security and Trust in Fog Network: Mecanismos de autenticação e confiança em nós de fog aumentam a latência

Desafios e Problemas em Aberto

- **Computing Task Related Challenges**
 - Big Data Analytics: Necessidade de um modelo capaz de gerar resultados corretos e capaz de aprender de fontes heterogêneas.
 - Data Fragmentation for Parallel Computing Exploitation: Busca pelo equilíbrio entre desempenho e tempo de treinamento, requer encontrar otimização para o número subconjuntos de dados.

Conclusão

- **Suporte a IoT:** Fog Computing integra recursos computacionais para aliviar a sobrecarga na nuvem e melhorar o desempenho do sistema.
- **Ambiente Complexo:** Heterogeneidade, mobilidade e mudanças dinâmicas dificultam a alocação eficiente de recursos.
- **Diversidade de Tarefas:** Diferentes características e demandas exigem soluções flexíveis.
- **Papel do RL:** Utilização para otimizar decisões em compartilhamento, escalonamento e offloading de tarefas, apesar dos desafios e incertezas.
- **Limitações dos Algoritmos Atuais:** Muitos algoritmos heurísticos não conseguem se adaptar de forma eficaz às incertezas da camada de Fog.
- **Desafios e Questões em Aberto:** Necessidade de explorar e solucionar diversos pontos a fim de encontrar uma solução mais eficiente.

Obrigado!

Av. Sete de Setembro, 3165 Rebouças
80230-901 Curitiba PR Brasil
+55 (41) 3310-4545

