



UNIVERSIDADE FEDERAL DO PARÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
DIRETORIA DE PESQUISA

RELATÓRIO TÉCNICO - CIENTÍFICO PARCIAL

PROGRAMA INSTITUCIONAL DE BOLSAS DE INICIAÇÃO CIENTÍFICA

PIBIC/CNPq, PIBIC/CNPq-AF, PIBITI, PIBIC/UFPA, PIBIC/UFPA-AF, PIBIC/UFPA-INTERIOR, e PIBIC/UFPA-EBTT, PIBIC/UFPA-PcD, PRODOUTOR, PRODOUTOR RENOVACÃO, PIVIC, FAPESPA, PIBIC-EM.

PERÍODO: _____/2020 a _____/2021

IDENTIFICAÇÃO DO PROJETO

Título do Projeto de Pesquisa (ao qual o Plano de Trabalho está vinculado): OTIMIZAÇÃO DO POSICIONAMENTO DE MÚLTIPLAS SMALL CELLS E UAVS USANDO APRENDIZADO DE MÁQUINA. Otimização do posicionamento de múltiplas small cells e UAVs usando aprendizado de máquina

Nome do Orientador: Jasmine Priscyla Leite de Araújo

Titulação do Orientador: Doutorado

Unidade (Campi/Instituto/Núcleo): Universidade Federal do Pará

Laboratório: Laboratório de Computação e Telecomunicações

Título do Plano de Trabalho: Otimização do posicionamento de UAVs como estações aéreas voadoras para atender alto tráfego com mobilidade.

Nome do Bolsista: Samuel Figueira Aguiar

Tipo de Bolsa:

- ☐ PIBIC/CNPq
- ☐ PIBIC/CNPq-AF
- ☐ PIBITI
- ☒ PIBIC/UFPA
- ☐ PIBIC/UFPA-AF
- ☐ PIBIC/UFPA-INTERIOR
- ☐ PIBIC/UFPA-EBTT
- ☐ PIBIC/UFPA-PcD
- ☐ PIBIC PRODOUTOR
- ☐ PIBIC PRODOUTOR RENOVACÃO
- ☐ PIVIC
- ☐ FAPESPA
- ☐ PIBIC-EM

- O relatório não deverá ultrapassar 10 MB ou conter mais de vinte (20) páginas.

1. Atividades realizadas:

Primeiramente, foram feitas algumas reuniões entre o bolsista e o orientador com a finalidade de explicar bem os objetivos da pesquisa. Esses encontros foram feitos de maneira online, por conta da pandemia de COVID-19, através da plataforma Google Meet, onde a professora Jasmine explicou detalhadamente como o trabalho seria executado.

Após isso, alguns artigos e livros foram selecionados para se verificar o que já havia sido produzido relacionado ao tema abordado. Artigos sobre movimentação e distribuição de usuários e redes utilizando drones foram separados. Destaque para os artigos “*Distributed Drone Base Station Positioning for Emergency Cellular Networks Using Reinforcement Learning*” e “*Coverage Optimization with a Dynamic Network of Drone Relays*”.

Esses artigos estão sendo lidos e resumidos. As informações úteis serão coletadas para a aplicação nas experimentações via simulação e vida real.

A proposta do trabalho pode ser resumida da seguinte forma: um estudo de técnicas de aprendizado de máquina para a localização ótima de *small cells* e *VANTs* no planejamento de monitoramento dos novos sistemas celulares.

A técnica de aprendizado de máquina utilizada será o aprendizado por reforço, mais especificamente a aprendizagem Q (Q-learning). O aprendizado por reforço é uma técnica que se baseia em um sistema de recompensas. Um agente é colocado em um ambiente e tem de aprender a se comportar com sucesso nesse ambiente [Inteligência Artificial, Stuart Russell, terceira edição]. Conforme ele explora o ambiente, ele é punido por errar ou se afastar de seu objetivo ou recompensado por se aproximar ou atingir a ele. Isso se repete inúmeras vezes, até o agente estar satisfatoriamente treinado para lidar com a atividade desejada.

Falando mais especificamente sobre o q-learning, diferente do aprendizado por reforço tradicional, o agente não necessita de um modelo do ambiente em que vai trabalhar. Isso ocorre porque na aprendizagem Q, o agente irá tomar decisões baseado na utilidade de suas ações em determinado estado, ao invés de verificar diretamente os resultados. Isso abre espaço para maior flexibilidade em relação aos ambientes que o agente pode atuar. Em contrapartida, os agentes de aprendizagem Q, por não saberem aonde suas ações levam, não podem realizar uma observação antecipada.

No projeto de pesquisa, por conta do drone precisar se adaptar a ambientes extremamente diferentes, o Q-Learning é indicado. A variedade se dá porque certamente o drone não atuará somente no mesmo espaço geográfico e muito menos com a mesma quantidade e localização de usuários utilizando a rede.

Para simularmos adequadamente usuários em movimento, é necessário o estudo do próprio movimento dos usuários. Para isso, estuda-se modelos de movimento. Modelos de movimento caracterizam os movimentos dos usuários móveis em relação à sua localização, velocidade e direção ao longo de um período. Eles são cruciais porque simulações sem uma movimentação realista produzem resultados irrealistas, levando a conclusões erradas.

De acordo com Jagdeep Singh et al., existem quatro categorias de modelos de movimento: baseados em rastreamento, estocásticos, sintéticos e baseados em mapas.

Modelos baseados em rastreamento são feitos através do rastreamento de movimentações reais. Esses modelos possuem um alto grau de realismo, mas são muito difíceis de implementar, complexos e não abrem muito espaço para a diversificação.

Modelos estocásticos, por outro lado, são modelos em que o movimento acontece de forma aleatória e controlada. Esses modelos já apresentam um extremo oposto em relação ao rastreamento,

pois são fáceis de implementar, simples e abrem muito espaço para a diversificação, mas não apresentam muito realismo. Nessa classificação temos modelos como *Random Walk*, *Random Direction* e *Random Waypoint*.

Modelos sintéticos são modelos matemáticos que capturam o movimento dos nós utilizando de obstáculos, caminhos e outros empecilhos. Eles não se movem de forma aleatória, mas capturam movimentos realistas. Dessa forma, encontramos um equilíbrio entre diversificação e realismo, pois existirão diversos episódios diferentes durante a simulação ao mesmo tempo que tais episódios são realistas. Nessa classificação encontramos os Baseados em Dependência Temporal, Baseados em Dependência Espacial e Baseados em Restrições Geográficas,

Finalmente, os modelos baseados em mapas são modelos feitos em mapas reais. Esse tipo de modelo abre espaço para a solução de problemas específicos, mas fecha quase completamente para a diversificação. Nessa classificação encontramos os baseados em rotas, tráfego humano, movimento de dia de trabalho, mobilidade de uma comunidade, caminho mais curto, além dos modelos de redes sociais.

Existem inúmeros modelos dentro dessas 4 categorias, mas dois foram selecionados inicialmente. Os dois modelos de movimentos selecionados serão o Manhattan Grid e o Random Walk, o primeiro sendo um modelo geográfico (que faz parte dos sintéticos) e o segundo do estocástico. Esses dois são os modelos de movimento mais comuns e utilizados em simulações e produzem resultados mais “brutos”, ou seja, de situações menos específicas.

Random Walk faz com que os usuários da simulação se movam de maneira completamente aleatória, em uma certa velocidade, em um espaço limitado pelo simulador. Dessa forma, pode-se treinar o algoritmo para lidar com situações de aleatoriedade em um campo aberto. Situações que podem se assemelhar com o Random Walk incluem animais com dispositivos de IOT em uma fazenda, pessoas em um grande concerto em espaços abertos etc.

Já o modelo Manhattan Grid trata-se de uma movimentação em áreas com obstáculos, mais especificamente em formato de grade. Esse modelo serve para simular ambientes urbanos, já que os usuários se movem na horizontal e na vertical, não entrando nos obstáculos (que seriam os prédios). Apesar de nem todos os ambientes urbanos se disporem em formato de grade, esse modelo de movimento é ótimo para tratar de casos em cidades.

O algoritmo proposto para a solução do problema é semelhante ao proposto no artigo de Klaine. O algoritmo procurará o melhor posicionamento para os múltiplos drones. O melhor posicionamento é aquele que conseguir maximizar a quantidade de usuários conectados à rede com uma conectividade decente.

Os agentes, cada um com sua própria matriz Q , irão explorar o ambiente. Eles correspondem aos drones com sinal.

Os estados são o posicionamento dos drones, seja em um ambiente 2D ou 3D. Como o posicionamento é uma matriz, seus posicionamentos são finitos.

As ações que podem ser tomadas são as movimentações possíveis, incluindo manter-se na posição ou mover-se na diagonal. Em um ambiente 3D também inclui subir ou descer.

A recompensa baseia-se no total de usuários alocados na rede construída por todos os drones. Não baseia-se em quantos usuários cada drone aloca, pois isso faria com que diferentes drones alocassem os mesmos usuários.

A inicialização se dá com uma matriz zerada e com os drones em posições aleatórias.

Os critérios de parada são baseados em três condições. Os drones se moveram por um número máximo de iterações estipulado, o valor da recompensa não melhorou em um certo número de iterações estipuladas, ou o drone usou todos os seus blocos de recursos e explorou por um número mínimo de iterações estipuladas.

O algoritmo pode ser resumido da seguinte forma:

- 1) Inicializar localização dos drones;
- 2) Inicializar matrizes Q ;
- 3) Para cada episódio, faça:

- 4) Enquanto o critério de parada não for atingido, faça:
- 5) Drones escolhem a ação com o maior valor de Q;
- 6) Alocar usuários;
- 7) Observar a recompensa, $r(t+1)$;
- 8) Atualizar o estado;
- 9) Atualizar matrizes Q;
- 10) Fim do faça;
- 11) Mover-se para a melhor recompensa observada;
- 12) Alocar usuários;
- 13) Gravar métricas;
- 14) Fim do faça;

2. Comparação entre o plano original e o executado:

(Relatar as tarefas desenvolvidas e se o plano de trabalho original foi executado ou passou por modificações).

Por enquanto, não houve mudanças significativas no trabalho além do deslocamento das atividades.

3. Outras atividades:

(Descrever trabalhos complementares não relacionados, especificamente, ao plano de trabalho).

Não houve trabalhos complementares não relacionados, especificamente, ao plano de trabalho.

4. Resultados preliminares

(Listar os resultados preliminares, de acordo com os objetivos e a metodologia constantes no plano original, bem como informar se houve modificações nestes itens).

5. Atividades desenvolvidas nos próximos meses:

Após a leitura e resumo dos artigos, um simulador será desenvolvido. A simulação será feita através de um programa que será desenvolvido em Python que leva em consideração várias variáveis de ambiente importantes para uma aproximação de uma situação real.

Um artigo deve ser escrito com os resultados obtidos com as simulações e submetido a uma conferência nacional.

6. Dificuldades:

(Relacionar os principais fatores negativos que interferiram na execução do plano de trabalho).

Certamente o agravamento da pandemia causou diversos atrasos no desenvolvimento do trabalho. O bolsista e a orientadora ficaram doentes, assim como suas respectivas famílias. Isso impediu com que o trabalho fosse feito e que reuniões acontecessem. Somente no início do ano de 2022 foi possível ter uma primeira reunião presencial, essa mais produtiva.

7. Referências Bibliográficas:

- NORVIG, Peter; RUSSELL, Stuart. Inteligência artificial: Tradução da 3a Edição. 3. ed. [S. l.]: Elsevier Brasil, 2014. 1056 p. ISBN 9788535251418.
- VERMA, Anshul et al. Opportunistic Networks: Fundamentals, Applications and Emerging Trends. 1. ed. [S. l.]: CRC Press, 2021. 330 p. ISBN 9780367677305.
- Klaine, P.V., Nadas, J.P.B., Souza, R.D. et al. Distributed Drone Base Station Positioning for Emergency Cellular Networks Using Reinforcement Learning. *Cogn Comput* 10, 790–804 (2018). <https://doi.org/10.1007/s12559-018-9559-8>
- E. Arribas, V. Mancuso and V. Cholvi, "Coverage Optimization with a Dynamic Network of Drone Relays," in *IEEE Transactions on Mobile Computing*, vol. 19, no. 10, pp. 2278-2298, 1 Oct. 2020, doi: 10.1109/TMC.2019.2927335.

PARECER DO ORIENTADOR: Manifestação do orientador sobre o desenvolvimento das atividades do aluno (o parecer deve ser cadastrado no SIGAA).

DATA: ____/____/____

(Dispensado em caso de envio via SIGAA)

ASSINATURA DO ORIENTADOR

(Dispensado em caso de envio via SIGAA)

ASSINATURA DO ALUNO