

MFA-RNN: Uma Rede Neural Recorrente para Predição de Próximo Local de Visita com Base em Dados Esparsos

Cláudio G. S. Capanema¹, Fabrício A. Silva¹, Thais R. M. B. Silva¹

¹Universidade Federal de Viçosa (Campus UFV-Florestal)
Florestal – MG – Brasil

{claudio.capanema, fabricio.asilva, thais.braga}@ufv.br

Abstract. *Predicting the mobility of a user is an important task to enhance the effectiveness of mobile applications. In this work, we present the MFA-RNN (Multi-Factor Attention Recurrent Neural Network), a neural network that uses the Multi-Head Self-Attention technique to extract correlations under several features of the sequence of the visited places. The proposed model is able to predict the next place of visit considering multiple factors (user, location, time and type of the day) of each record of the sequence. Moreover, we propose a method to fill sparse data to enhance the performance of the solution. The obtained results indicate the effectiveness of the MFA-RNN model in relation to four known solutions of the literature.*

Resumo. *Prever a mobilidade de um usuário é uma tarefa importante para se elevar a efetividade de aplicações móveis. Neste trabalho, é apresentada a MFA-RNN (Multi-Factor Attention Recurrent Neural Network), uma rede neural recorrente que utiliza a técnica Multi-Head Self-Attention para extrair correlações sob diversos aspectos da sequência de locais visitados. O modelo é capaz de prever o próximo local de visita considerando múltiplos fatores (usuário, localização, tempo e tipo do dia) de cada registro da sequência. Além disso, é proposto um método para o preenchimento de dados esparsos para melhorar o desempenho da solução. Os resultados obtidos indicam a eficácia do modelo MFA-RNN em relação a quatro soluções conhecidas na literatura.*

1. Introdução

A popularização dos dispositivos móveis trouxe consigo o conceito de *Location-Based Services (LBS)*, que contempla um conjunto de serviços destinados a tornar mais eficiente o dia a dia de usuários móveis [Al-Molegi and Martínez-Ballesté 2018]. Esses serviços têm como principal recurso a utilização de informação georreferenciada gerada pelos sensores dos dispositivos móveis. Considerando que o comportamento humano tende a seguir padrões de mobilidade de pessoa para pessoa [Banovic et al. 2016], é importante que as rotinas de milhares de usuários sejam estudadas.

Neste contexto, surge o problema de predição de próximo local de visita considerando o histórico de movimentação de um usuário. Essa tarefa é especialmente importante para *LBS* que necessitam de informação sobre o futuro próximo. Dentre as aplicações estão sistemas de recomendação [Wei et al. 2012], previsão de tráfego [Gao et al. 2016], melhoria na interação humana [Yao et al. 2016], dentre outras. Assim, ao se prever o

próximo local de visita, provedores de serviços podem aprimorar os seus produtos, e aumentar o engajamento e a satisfação de seus clientes.

Diversos métodos têm sido apresentados para se prever o próximo local de visita. As redes neurais recorrentes têm se destacado por tornar possível identificar padrões de mobilidade de diferentes agentes sob vários fatores. Recentemente, estudos têm sido conduzidos para avaliar quais tipos de informações trazem maior impacto na predição do próximo local de visita. No entanto, um aspecto pouco abordado, mas muito relevante, diz respeito ao método de coleta e o volume dos dados utilizados para avaliar as soluções desenvolvidas atualmente. Dados gerados por meio de *check-ins* de localização em redes sociais não são capazes de representar por completo a rotina humana, uma vez que é necessário que o usuário sempre informe a sua posição de forma ativa (i.e., coleta ativa). Como alternativa, alguns trabalhos recorrem a dados de *GPS* coletados entre pequenos intervalos de tempo sem a necessidade de interferência humana (i.e., coleta passiva), a fim de avaliar melhor as suas soluções. Esse método de coleta de informações, no entanto, dificilmente pode ser aplicado em larga escala, onde a elevada carga de utilização do sensor de *GPS* pode comprometer a autonomia energética dos dispositivos móveis. Assim, soluções desenvolvidas sob dados de *GPS* esparsos, onde podem existir longos períodos entre a coleta de informações, tendem a ser mais atrativas para serem aplicadas na prática.

Neste trabalho, é proposta a *MFA-RNN* (*Multi-Factor Attention Recurrent Neural Network*), uma rede neural recorrente para a predição de próximo local de visita baseada em dados esparsos. Comparando-se com estudos anteriores, a *MFA-RNN* engloba convenientemente o mecanismo de *Multi-Head Self-Attention* [Vaswani et al. 2017] com diferentes tipos de informações que podem ser obtidas dos traços de localização, e que não foram utilizadas simultaneamente em outros trabalhos. Além disso, é proposto um novo método de preenchimento de dados de *GPS* esparsos, a fim de reduzir os efeitos da falta de dados e elevar o desempenho da rede neural. Dessa forma, as contribuições deste trabalho são resumidas em:

- Proposta de uma nova arquitetura capaz de aplicar a técnica de *Multi-Head Self-Attention* no contexto de predição do próximo local de visita ao mesmo tempo que engloba diferentes fatores.
- Utilização de múltiplos fatores de entrada para o modelo: localização, horário, ID do usuário e o tipo do dia (dia de semana ou final de semana).
- Considerando o trabalho de [Capanema et al. 2019], é possível inferir se uma pessoa está em casa em horários específicos mesmo que não tenham sido gerados registros. Essa contribuição é estendida neste trabalho para preencher a base de dados esparsos, e assim, facilitar o reconhecimento de padrões pela rede neural.
- Diferentemente de outros trabalhos, é apresentada uma análise que indica como a rotina humana varia entre dias de semana e finais de semana. Juntamente com os resultados obtidos, a utilização da informação de tipo do dia do registro é, dessa forma, fundamentada tanto do ponto de vista teórico quanto prático.
- Validação da proposta *MFA-RNN* considerando-se dados de 5.272 usuários reais, o que representa a maior base de dados coletados de forma passiva entre os trabalhos existentes encontrados na literatura.

O restante do trabalho está organizado da seguinte forma. Inicialmente, a Seção 2 apresenta os principais trabalhos relacionados. A Seção 3 descreve a base de dados

e o método utilizado de preenchimento de dados esparsos. Na Seção 4, é apresentada a análise sobre a rotina dos usuários. Posteriormente, a arquitetura *MFA-RNN* é apresentada na Seção 5. Por último, os resultados dos experimentos, e a conclusão e trabalhos futuros estão organizados nas Seções 6 e 7, respectivamente.

2. Trabalhos Relacionados

Diversas técnicas já foram desenvolvidas para a predição de próximo local de visita. As redes neurais recorrentes são uma variação de rede neural que têm ganhado destaque sobre métodos tradicionais, como cadeias de *Markov* [Zeng et al. 2019], para solucionar o problema. A evolução das soluções tem ocorrido em dois principais aspectos: utilização de diferentes tipos de informações que compõem os traços de localização, e aprimoramento da arquitetura da rede neural com a implementação de novas camadas.

Os modelos *STF-RNN* [Al-Molegi et al. 2016] e *MAP* [Al-Molegi et al. 2018] consideram como entradas para seus modelos apenas informações espaço-temporais (localização e horário) dos eventos gerados pelos dispositivos móveis. Em geral, cada localização corresponde ao identificador do ponto de interesse (PoI) ou região de interesse (RoI) visitado pelo usuário. Já a informação temporal é comumente representada pela hora do dia em que o evento foi gerado. Em busca de aprimoramentos, [Al-Molegi and Martínez-Ballesté 2018] concluíram, através do modelo chamado *STW-RNN + WKD/WKE*, que o tipo do dia (dia de semana ou final de semana) do evento é uma informação relevante para a rede neural. No entanto, não foram conduzidas análises do quanto a rotina humana pode variar ao longo da semana. Por outro lado, os trabalhos [Feng et al. 2018, Yao et al. 2017, Zeng et al. 2019] avançaram na área ao incorporarem à rede neural o identificador do usuário que gerou cada evento, tornando possível que o modelo aprenda a rotina específica de cada pessoa. Além disso, na proposta de [Yao et al. 2017], de nome *SERM*, foi introduzida a utilização de informações textuais de dados do *Twitter* e *Foursquare* através de uma camada *Embedding* associada com a técnica *GloVe* [Pennington et al. 2014].

Com relação à arquitetura de redes neurais, os autores de [Vaswani et al. 2017] introduziram uma nova camada de nome *Multi-Head Self-Attention (MHSA)*, originalmente desenvolvida para o problema de Tradução de Máquina. O mecanismo é capaz de extrair correlações sob diferentes partes de uma sequência, e assim como demonstrado em [Zeng et al. 2019], pode ser aplicado no problema de predição de próximo local de visita. A solução de [Zeng et al. 2019] é referida como *MHSA+PE* no presente trabalho, onde *PE (Positional Encoding)* é uma camada utilizada para propagar a noção de ordem dos eventos de uma sequência. Já em [Al-Molegi et al. 2018], um mecanismo de atenção simplificado foi utilizado no modelo *MAP*. Dentre as camadas recorrentes, as soluções de [Al-Molegi et al. 2016, Al-Molegi et al. 2018] utilizaram a *Simple RNN*, enquanto em [Yao et al. 2017] a *Long Short-Term Memory (LSTM)* foi aplicada. Os autores de [Zeng et al. 2019, Feng et al. 2018] recorreram a uma camada *Gated Recurrent Unit (GRU)* que, assim como a *LSTM*, tem o diferencial de aprender melhor com longas sequências.

Neste trabalho, a solução proposta *MFA-RNN* engloba as principais técnicas descritas na literatura em termos de redes neurais (camadas *MHSA* e *GRU*, e múltiplas entradas), para a predição do próximo local de visita. Além disso, diferentemente das soluções

existentes, a proposta também considera que os dados são esparsos e portanto, realiza um preenchimento dos mesmos. A Tabela 1 sumariza as principais soluções que envolvem redes neurais, e as compara com o presente trabalho. Em geral, as soluções da literatura são avaliadas sob um grande número de usuários quando o método de coleta de dados é ativo (e.g., *check-ins*). Porém, quando o método de coleta é passivo, poucos usuários de teste são considerados. Para aferir a qualidade da proposta *MFA-RNN*, as soluções *MAP* [Al-Molegi et al. 2018], *STF-RNN* [Al-Molegi et al. 2016], *SERM* [Yao et al. 2017] e *MHSA+PE* [Zeng et al. 2019] são utilizadas como base de comparação neste trabalho.

Tabela 1. Comparação com trabalhos da literatura.

Solução	Camada recorrente	Dados de entrada	Mecanismo de atenção	Usuários avaliados
<i>MAP</i>	<i>Simple RNN</i>	Localização e tempo	Simples	182 (coleta passiva), 319.063 (coleta ativa)
<i>STF-RNN</i>	<i>Simple RNN</i>	Localização e tempo	Não possui	182 (coleta passiva)
<i>SERM</i>	<i>LSTM</i>	Localização, tempo, ID do usuário e informação contextual	Não possui	7.826 (coleta ativa)
<i>DeepMove</i>	<i>GRU</i>	Localização, tempo e ID do usuário	Simples	15.639 (coleta ativa), 1.075 (coleta passiva)
<i>STW-RNN+WKD/WKE</i>	<i>Simple RNN</i>	Localização, tempo e tipo do dia	Não possui	182 (coleta passiva), 319.063 (coleta ativa)
<i>MHSA+PE</i>	<i>GRU</i>	Localização, tempo e ID do usuário	<i>MHSA</i>	1.083 (coleta ativa)
<i>MFA-RNN (Proposta Atual)</i>	<i>GRU</i>	Localização, tempo, ID do usuário e tipo do dia	<i>MHSA</i>	5.272 (coleta passiva)

3. Descrição dos Dados

3.1. Características

O conjunto de dados utilizado foi fornecido sob confidencialidade por um provedor de serviços privado. Foram coletados, voluntariamente, dados de 5.272 usuários de dispositivos móveis durante um período de 62 dias. Cada evento de localização contendo a coordenada geográfica, o horário e o identificador do usuário era gerado quando o usuário se deslocava em pelo menos 100 metros, e desbloqueava a tela do seu aparelho. Apesar da necessidade de o usuário desbloquear a tela para que a coleta fosse realizada, diferentemente de *check-ins* utilizados na coleta ativa, o desbloqueio da tela é uma ação usual feita constantemente pelos usuários, principalmente em seus pontos de interesse. Esse método de coleta leva a uma base de dados esparsa, em que longos períodos podem existir entre a coleta de um evento e outro. As Figuras 1 e 2 exibem gráficos de CDF (*Função de distribuição acumulada*), obtidos pela distância e o tempo entre eventos consecutivos. Existe 50% de probabilidade de que a distância entre eventos consecutivos seja maior do que 1.049 metros e 50% de probabilidade para o tempo seja maior do que 35 minutos.

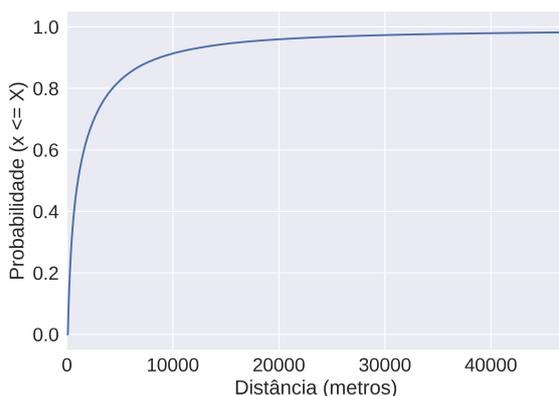


Figura 1. CDF de distância entre cada par de registros consecutivos.

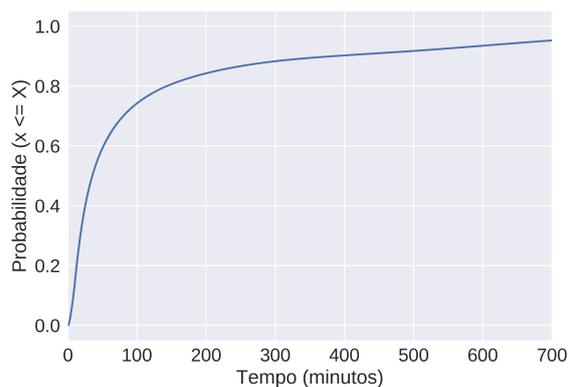


Figura 2. CDF do intervalo de tempo entre cada par de registros consecutivos.

3.2. Preenchimento

Para cada usuário, é necessário primeiro definir a sua mobilidade em termos de pontos de interesse visitados. Para isso, foi utilizada a solução proposta em [Capanema et al. 2019], desenvolvida sob dados esparsos. Foram identificados e classificados dois tipos de pontos de interesses pessoais: *Casa* e *Outro*. A detecção do PoI *Trabalho* possui uma assertividade baixa, e dessa forma, optou-se considerá-lo como *Outro* neste trabalho. Assim, para cada usuário é possível identificar o seu local de moradia e PoIs diversos que são definidos como *Outro*. Além disso, um terceiro estado de localização é considerado quando o usuário não está em nenhum PoI: *Deslocamento*. Essas informações são utilizadas para construir o conjunto de sequências de locais visitados para cada usuário. Cada sequência $S(u)$ de um usuário u é composta por N eventos $e = (l, t, id, td)$ gerados durante a coleta de dados. Os elementos de um evento correspondem, respectivamente, ao tipo da localização (*Casa*, *Outro* ou *Deslocamento*), hora do dia, usuário e tipo do dia (dia de semana e final de semana). O problema de aplicar puramente essa abordagem, no entanto,

é a existência de longos períodos entre um evento e outro, o que reduz a chance de real correlação entre a ordem dos locais visitados por uma pessoa.

Considerando o método de coleta descrito anteriormente, o trabalho de [Capanema et al. 2019] mostrou que o usuário tende a estar em casa durante o maior período de tempo onde não foram gerados eventos por um usuário. Isso ocorre porque é necessária a ação do usuário com o desbloqueio do aparelho e uma movimentação mínima para se gerar dados de localização. Porém, no trabalho anterior, essa informação havia sido utilizada apenas para detectar PoIs do tipo Casa, e não para preencher sequências. Dessa forma, neste trabalho, esse conhecimento é estendido e aplicado para preencher a sequência de localizações do usuário com o PoI *Casa* durante o período de inatividade do usuário.

4. Análise de Rotina

A mobilidade humana pode variar de acordo com diversos fatores, dentre eles os tipos dos dias de uma semana (dia de semana ou final de semana). Nesta seção, é apresentado um breve estudo que objetiva analisar o quanto a rotina humana pode variar ao longo da semana, e conseqüentemente, justificar o aproveitamento dessa informação na proposta *MFA-RNN*.

Após a identificação e classificação dos PoIs de um usuário, é possível caracterizar a sua rotina em termos da sua localização para cada hora do dia. Para isso, são contabilizadas as quantidades de eventos gerados em cada tipo de localização (*Casa*, *Outro* e *Deslocamento*) de acordo com cada hora do dia. Dessa forma, é possível estabelecer as probabilidades de se estar em *Casa* às 20:00 horas ou em *Deslocamento* às 11:00 horas, por exemplo.

Tabela 2. Exemplo da rotina de um usuário.

Horário	00:00-01:00	01:00-02:00	...	22:00-23:00	23:00-00:00
Casa	0.33	0.33	...	0.70	0.80
Outro	0.33	0.33	...	0.10	0.10
Deslocamento	0.33	0.33	...	0.20	0.10
Entropia	1.10	1.10	...	0.80	0.64

A Tabela 2 exemplifica a distribuição de probabilidades para um usuário, em que para cada hora do dia é calculada a entropia de *Shannon* [Shannon 1948], com base e , para medir o nível de incerteza do respectivo conjunto de probabilidades. Assim, para cada hora se obtém o quão previsível é a rotina do usuário. Como o trabalho é avaliado sob dados esparsos, é comum que em algumas horas do dia nenhum registro tenha sido gerado por um determinado usuário. Nesse caso, as probabilidades de se estar em Casa, Outro e Deslocamento ser tornam iguais, e a entropia correspondente tem o valor máximo de 1.10, indicando elevada incerteza sobre a rotina naquele horário. A média das entropias obtidas em cada horário indica o quão previsível pode ser a rotina de uma determinada pessoa considerando o seu padrão de mobilidade. Intuitivamente, a rotina humana tende a variar ao longo da semana. Dessa forma, a caracterização de rotina é feita separadamente para dias de semana e finais de semana.

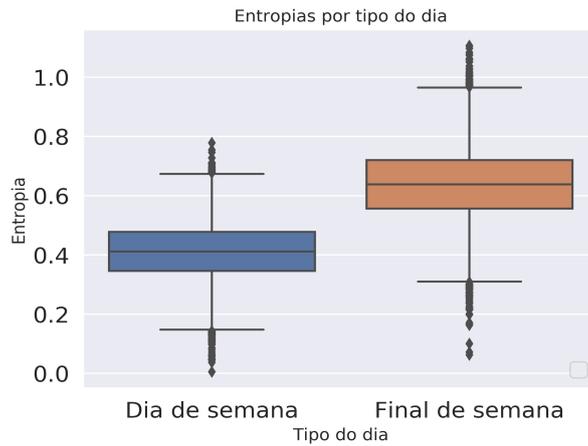


Figura 3. Distribuição das entropias médias de todos os usuários. Quanto menor o valor, mais previsível é a rotina.

A Figura 3 apresenta as distribuições das entropias de todos os usuários separadamente para dias de semana e finais de semana. Pode-se perceber que existe uma maior previsibilidade da rotina dos usuários analisados durante os dias de semana, uma vez que o valor da entropia é menor. Por outro lado, o valor maior de entropia durante finais de semana indica que as rotinas tendem a variar muito nesses dias.

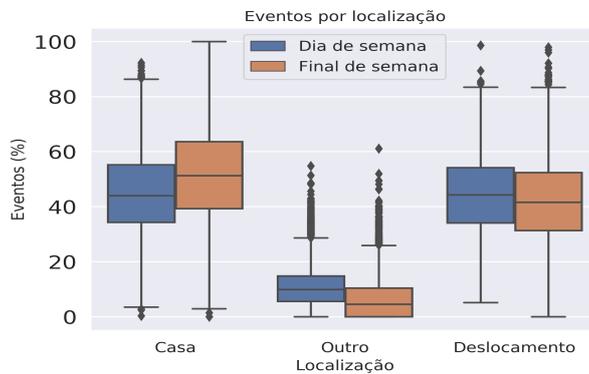


Figura 4. Comparação entre a porcentagem de eventos de cada localização gerados em dia de semana e final de semana.

Outra informação importante é que a proporção de eventos gerados em cada localização varia de acordo com o tipo de dia. A Figura 4 indica que existem proporcionalmente mais eventos em *Casa* nos finais de semana do que nos demais dias. Por outro lado, os pontos de interesse do tipo *Outro* têm um menor peso no número de visitas nos sábados e nos domingos do que nos dias de semana. Similarmente, os usuários têm uma tendência menor de estar em *Deslocamento* durante os finais de semana, em relação aos demais dias.

Portanto, indicar o dia da semana em que um evento foi gerado tende a ser uma informação relevante para uma rede neural prever o próximo ponto de visita de um usuário. A variação da entropia, e a troca de proporção de eventos entre *Casa* e *Outro* durante a semana, sugerem as seguintes hipóteses:

1. Existe uma probabilidade maior para que o próximo local de visita seja *Casa* nos finais de semana do que nos demais dias. Por outro lado, de segunda-feira à sexta-feira os usuários tendem a estar em *Deslocamento* e visitar mais o PoI *Outro*.
2. Em geral, durante dias de semana, menos localizações diferentes são visitadas em um mesmo horário, o que reduz a entropia. Assim, a tendência é que cada horário esteja associado a uma localização predominante apenas.

Assim, a utilização pela rede neural da informação de tipo de dia da semana em que eventos foram gerados pode ser fundamentada do ponto de vista teórico.

5. Arquitetura MFA-RNN

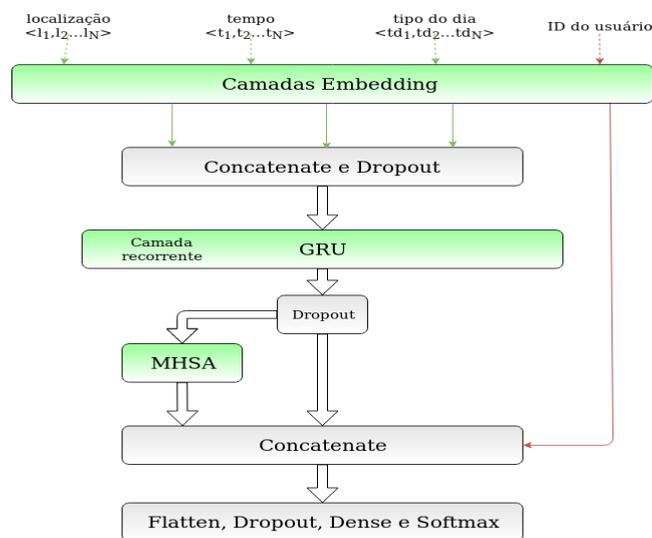


Figura 5. Arquitetura MFA-RNN.

Redes neurais recorrentes são tradicionalmente utilizadas para processar sequências de informações textuais, onde se busca compreender, por exemplo, as correlações entre a ordem de ocorrência de palavras de uma frase. Em uma interpretação mais ampla, as palavras podem ser compreendidas como variáveis categóricas, que no contexto deste trabalho correspondem aos elementos que compõem cada evento (ver Seção 3.2).

A rede neural proposta MFA-RNN, ilustrada na Figura 5, é capaz de aprender sob múltiplos fatores (localização, tempo, usuário, tipo do dia) para prever o próximo local de visita. Diversas camadas (*Embedding*, *Concatenate*, *GRU*, *MHA* e *Dropout*) estão convenientemente dispostas, a fim de se alcançar um modelo eficiente e generalista.

A seguir, a arquitetura da rede neural é descrita destacando-se as características das principais camadas utilizadas.

5.1. Camadas *Embedding*

A entrada da rede neural pode ser compreendida como uma composição de quatro matrizes de localização, tempo, tipo do dia e ID do usuário, que representam os eventos de uma sequência codificados em inteiros positivos. A função das camadas *Embeddings* é representar cada um desses inteiros na forma de vetores numéricos de tamanho fixo. Assim, os

valores de cada representação vetorial são atualizados durante o treinamento a fim de que palavras que tendem a ocorrer próximas nas sequências tenham representações vetoriais parecidas. Além disso, a utilização de vetores densos torna possível trabalhar eficientemente com grandes vocabulários de palavras, o que não ocorre com a técnica de *one-hot-encoding*, onde os vetores tendem a ser de grande dimensão e esparsos [Feng et al. 2018].

Após a camada *Embedding*, os vetores de saída de localização, tempo e tipo do dia são concatenados, e a técnica *Dropout* é aplicada com o intuito de evitar sobre-ajuste de treinamento e assim obter um modelo mais generalista.

5.2. Camada recorrente: *Gated Recurrent Unit (GRU)*

A intuição por trás de uma camada recorrente é que existe uma ordem lógica na disposição dos elementos em uma sequência, que portanto deve ser aprendida. No contexto deste trabalho, a ordem representa a sequência de locais visitados pelo usuário.

Diferentemente da *Simple RNN*, as camadas *LSTM (Long Short-Term Memory)* e *GRU (Gated Recurrent Unit)* são variantes de redes neurais recorrentes capazes de extrair correlações entre eventos de longas sequências [Fu et al. 2016]. A *GRU* engloba informações de passos anteriores e da entrada corrente, de onde são selecionadas as informações a serem processadas. Em termos práticos, a vantagem da *GRU* sobre a *LSTM* está no seu custo computacional [Fu et al. 2016], e por este motivo foi utilizada no modelo proposto *MFA-RNN*. Por fim, a técnica *Dropout* é aplicada para evitar sobre-ajuste.

5.3. *Multi-Head Self-Attention (MHSA)*

O mecanismo *Self-Attention* foi inicialmente desenvolvido para o problema de Tradução de Máquina, onde se buscava compreender a correlação entre as palavras de uma sequência em um determinado idioma, com palavras de uma segunda sequência em outro idioma. No contexto deste trabalho, ao invés de se comparar sequências diferentes, as palavras da própria sequência são correlacionadas, e por isso o nome dado *Self-Attention*. A técnica *Multi-Head Self-Attention* introduzida em [Vaswani et al. 2017], aplica o *Self-Attention* sobre diferentes partes do vetor de entrada da camada, sendo cada parte nomeada de *head*. A quantidade de *heads* que trouxe melhores resultados foi 4, indicando que existem 4 fatores que devem ser processados separadamente. Portanto, a partir da camada *MHSA* é possível “prestar atenção” sobre diferentes subespaços e fatores, correlacionando elementos de um mesmo vetor.

5.4. Englobando resultados

O mecanismo descrito na Seção 5.3 não consegue extrair correlações com base na ordem das sequências por si só [Vaswani et al. 2017]. Por esse motivo, a noção de ordem das sequências, obtida pela camada recorrente (*GRU*), é propagada diretamente. Dessa forma, as saídas das camadas *MHSA*, *GRU/Dropout* e *Embedding* para o ID do usuário, são concatenadas em um tensor resultante. Assim, a rede neural pode aprender sobre o perfil de mobilidade de um determinado usuário sob a perspectiva de duas técnicas simultâneas, *MHSA* e *GRU*. Em seguida, é realizada a redução de dimensão dos dados com *Flatten*, e o *Dropout* é aplicado para evitar sobre-ajuste. Por fim, uma camada *Dense* é utilizada juntamente com a função de ativação *softmax* para prever os três tipos de localização possíveis para o próximo local de visita.

6. Resultados e Análises

6.1. Configuração

Foram conduzidos experimentos com o objetivo de comparar o desempenho da rede neural proposta MFA-RNN, com trabalhos bem conhecidos da literatura. Como a base de dados é desbalanceada, ou seja, não possui uma quantidade semelhante de eventos das três classes de localizações possíveis (Casa, Outro e Deslocamento), a utilização da métrica de acurácia não é a mais adequada. Dessa forma, o objetivo de cada abordagem se torna obter o melhor desempenho para prever o próximo local de visita mediante as métricas Precisão, Revocação e F1-score. Cada modelo foi testado sob o método de validação cruzada, dividindo-se a trajetória de cada usuário em 5 partes.

As abordagens *MHSA+PE* ([Zeng et al. 2019]), *STF* ([Al-Molegi et al. 2016]), *SERM* ([Yao et al. 2017]) e *MAP* ([Al-Molegi et al. 2018]) são recentes e utilizaram uma grande variedade de técnicas em seus respectivos modelos. Portanto, elas foram selecionadas como bases de comparação com a solução proposta *MFA-RNN*.

Neste trabalho, foram conduzidos dois experimentos. No primeiro cenário, tanto o método proposto quanto as soluções base utilizam a base de dados esparsos. Para verificar o impacto positivo que o preenchimento dos dados traz, no segundo experimento a técnica de preenchimento (Ver Seção 3.2) é adotada por todas as soluções.

6.2. Avaliação Geral

O objetivo do primeiro experimento é avaliar o desempenho do modelo *MFA-RNN* e das *baselines* sob a base de dados original, ou seja, sem a aplicação do método proposto de preenchimento de dados. As Figuras 6, 7 e 8 exibem a Precisão, Revocação e *F1-score* de cada método, para os três tipos de localização existentes.

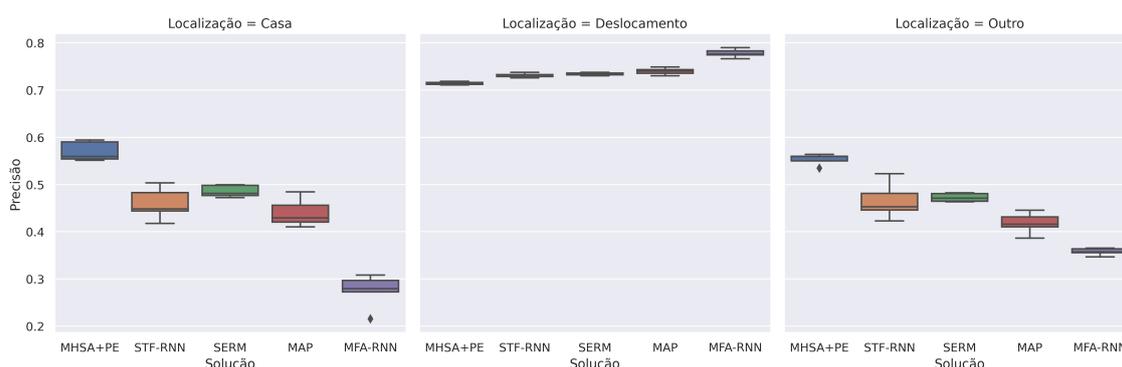


Figura 6. Precisão por localização (dados originais).

Apesar de o desempenho para se prever quando um usuário estará em *Deslocamento* ser ligeiramente inferior ao das soluções base, o *F1-score* da proposta *MFA-RNN* se manteve acima de 70%. Com relação à previsão dos PoIs *Casa* e *Outro*, as *baselines*, em geral, apresentam alta Precisão e baixa Revocação, o que afeta negativamente no valor final do *F1-score*.

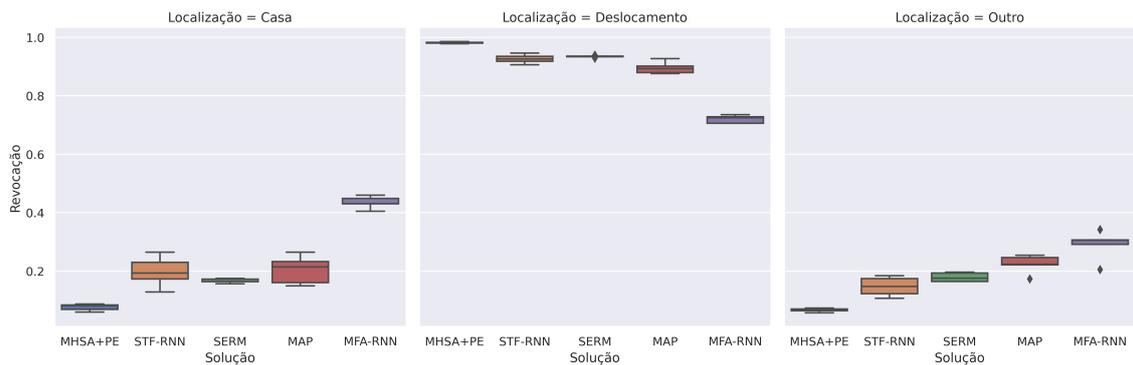


Figura 7. Revocação por localização (dados originais).

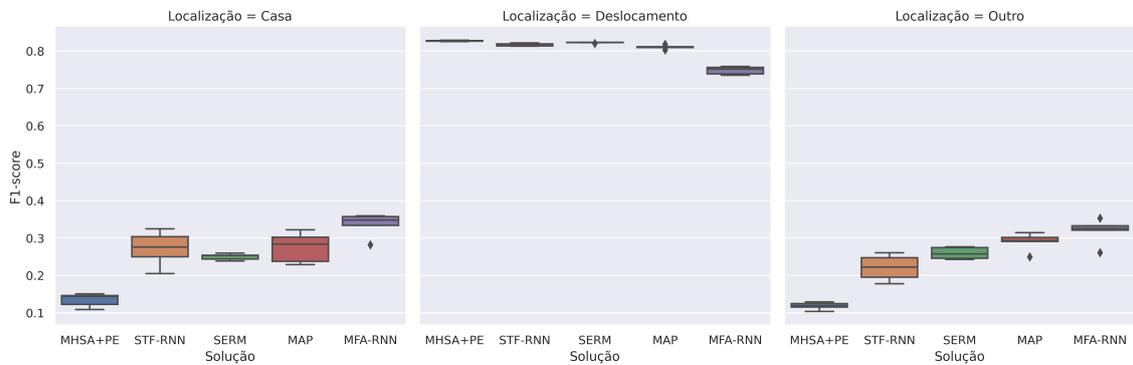


Figura 8. F1-score por localização (dados originais).

Ao mesmo tempo, os resultados indicam que o modelo *MFA-RNN* apresentou melhorias justamente onde as soluções base são deficitárias, ou seja, no *F1-score* de predição dos locais *Casa* e *Outro*. Dessa forma, é possível observar o ganho obtido com a arquitetura proposta, destacando-se a utilização da camada *MHSA*, e de múltiplos fatores como entrada pela rede neural, em especial através da *feature* “tipo do dia”.

6.3. Impacto do Preenchimento dos Dados

Neste experimento, todos os métodos são comparados sob a mesma base de dados preenchida. Assim, é possível avaliar o impacto do preenchimento de dados no desempenho das soluções. As Figuras 9, 10 e 11 exibem os desempenhos de cada método de acordo com as métricas Precisão, Revocação e *F1-score*, respectivamente.

O preenchimento de dados possibilitou que os desempenhos das soluções base alcançassem valores acima de 70% de *F1-score* para prever o PoI Casa. Juntamente com a melhoria na predição de Casa, as soluções base se aproximaram de 30% de *F1-score* para prever a localização do tipo Outro (exceto a solução *MHSA+PE*). Por outro lado, o desempenho para Deslocamento foi levemente reduzido entre essas abordagens, e, neste caso, o valor final de *F1-score* se aproximou da proposta *MFA-RNN*. Apesar de a solução *MHSA+PE* possuir a camada *MHSA*, a arquitetura apresentada não foi capaz de gerar bons resultados em ambos os cenários avaliados. A camada implementada *PE* (*Positional Encoding*) não contribuiu com a previsão do PoI Outro. A abordagem similar *MAP*

utilizou um mecanismo simples de *Attention*, e optou por propagar a noção de ordem dos elementos da sequência diretamente pela camada recorrente em detrimento da utilização de *PE*, o que trouxe melhores resultados. Essa intuição é aproveitada no presente trabalho, mas com o diferencial de se utilizar a técnica mais recente de *Multi-Head Self-Attention*.

Além disso, neste experimento o modelo *MFA-RNN* também se destaca quando o próximo local de visita é do tipo *Outro*, ao alcançar cerca de 35% de *F1-score*. A eficácia do método para *Deslocamento* é similar às soluções base, e portanto, mesmo aplicando preenchimento de dados em todas as abordagens, a rede neural proposta *MFA-RNN* apresenta um desempenho geral superior.

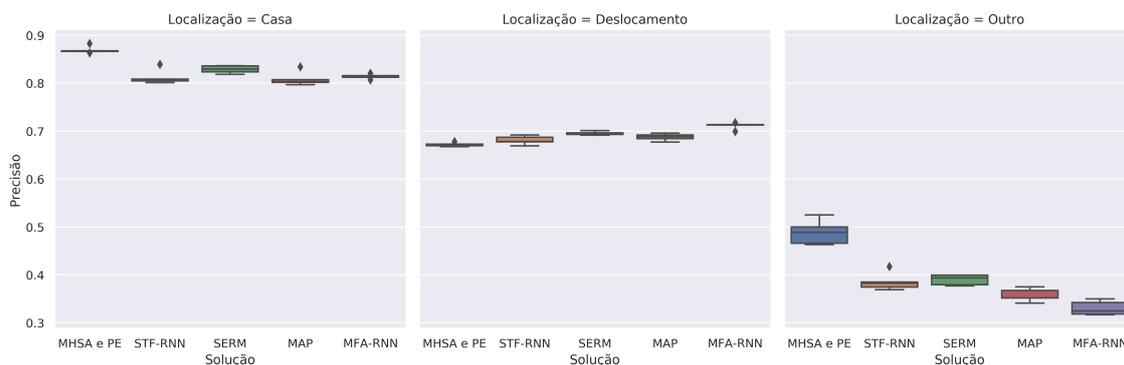


Figura 9. Precisão por localização (dados preenchidos).

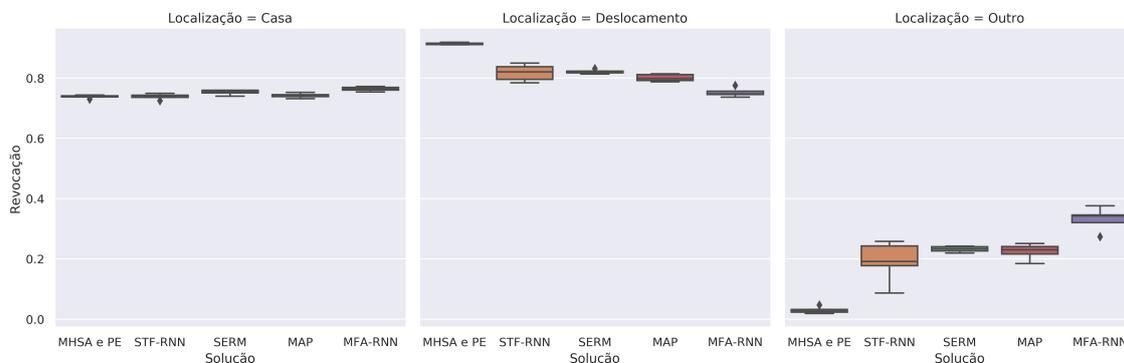


Figura 10. Revocação por localização (dados preenchidos).

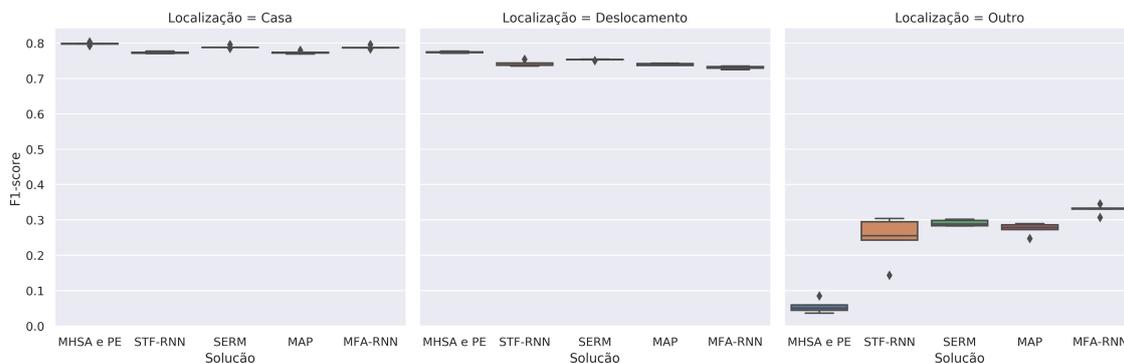


Figura 11. F1-score por localização (dados preenchidos).

Tanto no experimento anterior como no atual, a utilização de múltiplos dados de entrada gerou resultados mais consistentes (menor dispersão de valores de *F1-score*) para as soluções *MFA-RNN* e *SERM*. Comparando-se *SERM* com *STF-RNN* (cuja dispersão de *F1-score* é alta), as principais diferenças entre essas abordagens são que em *SERM* foi utilizada como entrada o ID do usuário, e o horário foi dividido em 48 partes (24 horas para dia de semana e 24 horas para finais de semana), funcionando analogamente à entrada de tipo do dia.

Por último, conclui-se que o método de preenchimento de dados pode contribuir com o desempenho tanto das soluções base quanto da rede neural apresentada. Em especial, ocorreram melhorias na predição dos PoIs Outro e, principalmente, Casa.

7. Conclusões e Trabalhos Futuros

Neste trabalho, foi apresentada a rede neural *MFA-RNN* para a predição de próximo local de visita de usuários móveis. O modelo se destaca por utilizar, ao mesmo tempo, múltiplos fatores de entrada da rede neural e o mecanismo de estado da arte *Multi-Head Self-Attention*. Com o objetivo de melhorar o desempenho do método desenvolvido, também foi apresentada uma abordagem para o preenchimento de dados esparsos, que é capaz de inferir quando um usuário esteve no PoI Casa. O modelo *MFA-RNN* obteve melhores resultados que as soluções base principalmente quando o próximo local de visita for do tipo Casa ou Outro. Os experimentos também demonstraram que o método de preenchimento de dados também pode ser utilizado para melhorar o desempenho de outras soluções da literatura. Além disso, nossos testes foram realizados a partir de uma base de dados contendo 5.272 usuários, superando os principais trabalhos da literatura que utilizam bases de dados esparsas, contendo dados georreferenciados de alta precisão e que foram coletados passivamente.

Como trabalhos futuros, poderão ser previstas mais categorias de PoI como Lazer, Hospedagem e Saúde, por exemplo, a fim de enriquecer os resultados. Dessa forma, o modelo *MFA-RNN* poderá prever próximos locais de visita de tipos mais específicos, o que ampliará ainda mais a sua efetividade.

Agradecimentos

Este trabalho contou com o apoio da CAPES, CNPq e Fapemig.

Referências

- Al-Molegi, A., Jabreel, M., and Ghaleb, B. (2016). Stf-rnn: Space time features-based recurrent neural network for predicting people next location. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–7. IEEE.
- Al-Molegi, A., Jabreel, M., and Martínez-Ballesté, A. (2018). Move, attend and predict: An attention-based neural model for people’s movement prediction. *Pattern Recognition Letters*, 112:34–40.
- Al-Molegi, A. and Martínez-Ballesté, A. (2018). The effect of space-time representation learning in predicting people’s next location. In *CCIA*, pages 64–73.
- Banovic, N., Buzali, T., Chevalier, F., Mankoff, J., and Dey, A. K. (2016). Modeling and understanding human routine behavior. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 248–260. ACM.

- Capanema, C. G. S., Silva, F. A., and Silva, T. R. M. B. (2019). Identificação e classificação de pontos de interesse individuais com base em dados esparsos. *Anais do XXXVII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos - SBRC 2019*, pages 16–29.
- Feng, J., Li, Y., Zhang, C., Sun, F., Meng, F., Guo, A., and Jin, D. (2018). Deepmove: Predicting human mobility with attentional recurrent networks. In *Proceedings of the 2018 World Wide Web Conference*, pages 1459–1468. International World Wide Web Conferences Steering Committee.
- Fu, R., Zhang, Z., and Li, L. (2016). Using lstm and gru neural network methods for traffic flow prediction. In *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pages 324–328. IEEE.
- Gao, J., Sun, Y., Liu, W., and Yang, S. (2016). Predicting traffic congestions with global signatures discovered by frequent pattern mining. In *2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pages 554–560. IEEE.
- Pennington, J., Socher, R., and Manning, C. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.
- Wei, L.-Y., Zheng, Y., and Peng, W.-C. (2012). Constructing popular routes from uncertain trajectories. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 195–203. ACM.
- Yao, D., Zhang, C., Huang, J., and Bi, J. (2017). Serm: A recurrent model for next location prediction in semantic trajectories. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 2411–2414. ACM.
- Yao, Z., Fu, Y., Liu, B., Liu, Y., and Xiong, H. (2016). Poi recommendation: A temporal matching between poi popularity and user regularity. In *2016 IEEE 16th international conference on data mining (ICDM)*, pages 549–558. IEEE.
- Zeng, J., He, X., Tang, H., and Wen, J. (2019). A next location predicting approach based on a recurrent neural network and self-attention. In *International Conference on Collaborative Computing: Networking, Applications and Worksharing*, pages 309–322. Springer.