



Douglas Diniz Landim

ANÁLISE DE DEMANDA VIA INTELIGÊNCIA ARTIFICIAL NO RESTAURANTE UNIVERSITÁRIO DO INSTITUTO CENTRAL DE TECNOLOGIA DA UNIFESP

São José dos Campos, SP

Douglas Diniz Landim

ANÁLISE DE DEMANDA VIA INTELIGÊNCIA ARTIFICIAL NO RESTAURANTE UNIVERSITÁRIO DO INSTITUTO CENTRAL DE TECNOLOGIA DA UNIFESP

Trabalho de conclusão de curso apresentado ao
Instituto de Ciência e Tecnologia – UNIFESP,
como parte das atividades para obtenção do tí-
tulo de Bacharel em Ciência da Computação.

Universidade Federal de São Paulo – UNIFESP

Instituto de Ciência de Tecnologia

Bacharelado em Ciência da Computação

Orientador: Prof. Dr. Vinicius Veloso

São José dos Campos, SP

Outubro de 2018

Douglas Diniz Landim

ANÁLISE DE DEMANDA VIA INTELIGÊNCIA ARTIFICIAL NO RESTAURANTE UNIVERSITÁRIO DO INSTITUTO CENTRAL DE TECNOLOGIA DA UNIFESP

Trabalho de conclusão de curso apresentado ao
Instituto de Ciência e Tecnologia – UNIFESP,
como parte das atividades para obtenção do tí-
tulo de Bacharel em Ciência da Computação.

Trabalho para apresentar em NOVEMBRO/2018:

Prof. Dr. Vinicius Veloso
Orientador

Professor
Convidado 1

Professor
Convidado 2

Professor
Convidado 3

São José dos Campos, SP
Outubro de 2018

Este trabalho é dedicado aos meus pais que apoiaram e sacrificaram esforços para me manter ativo nessa jornada, a todos os professores que me somaram conhecimentos, oportunidades e esperanças indo além de suas rotinas e agendas em prol do ensino, e principalmente à todos que me motivaram me oferecendo desafios para que eu pudesse enfrentá-los superando meus próprios limites

Agradecimentos

Minha jornada pela graduação foi marcada por muita persistência, dificuldades e fracassos. Agradeço primeiramente a Deus por me dar fé e alimentar minha persistência e esperança. Apesar de todo o conteúdo técnico das mais de 40 disciplinas do meu curso, o que mais me agregou evolução foi o ambiente extremamente desafiador desta universidade; que somado à muitas dificuldades pessoais, acidentes, contra-tempos de saúde, profissão e família; constituiu a soma perfeita de desafios que me derrubaram muitas e muitas vezes e me fizeram ser uma pessoa melhor, mais convicta e mais perseverante a cada nova tentativa de conquistar minhas aprovações. Agradeço à minha família por sempre me apoiar dando tudo de si, a meu professor por me aceitar orientar, e me motivar sempre me dando atenção como um amigo nas conversas no fim da aula durante o caminho até o estacionamento da faculdade, nas reuniões e chats online até nos finais de semana, aos amigos universitários, e a todos os professores que me acompanharam e ofereceram desafios em todos esses anos na UNIFESP.

*Mesmo desacreditado e ignorado por todos, não posso desistir, pois para mim, vencer é nunca
desistir.
(Albert Einstein)*

Resumo

O presente trabalho tem como objetivo a comparação com o método de análise da previsão de vendas do restaurante universitário da Unifesp, previamente feita pelo autor deste projeto com aplicação de métodos estatísticos, e atualmente com métodos de aprendizado de máquina. Tal análise por aprendizado de máquina foi apontado como relevante solução no fim da análise estatística do trabalho anterior. A temperatura da região onde se localiza o campus do restaurante foi analisada por recorrência via bootstrap como um fator que exerce impacto sobre as vendas do restaurante em certos períodos do semestre, e neste trabalho de conclusão de curso serão obtidas novas variáveis e um intervalo maior de amostragem na análise da previsão de demanda de refeições do restaurante em um novo modelo com aprendizado de máquina a fim de que seja obtido um modelo de previsão viável para evitar super-projeção de demanda com consequência de desperdício de alimentos, ou subprojeção com consequência de docentes ou discentes sem refeições.

Palavras-chaves: Redes Neurais Artificiais. Previsão de demanda. Aprendizado de Máquina. Inteligência Artificial. Perceptron Múltiplas camadas.

Abstract

The present work has the objective of comparison with the static method of analysis of the demand prediction in UNIFESP university restaurant, previously done by the author of this project, and currently with machine learning methods. Such analysis by machine learning was pointed out as a relevant solution at the end of the statistical analysis of the previous work. The temperature of the region where the restaurant campus is located was analyzed by bootstrap recurrence as a factor impacting the restaurant sales in certain periods of the semester, and in this work of completion of the course will be obtained new variables and a greater range of sampling in the analysis of the forecast of restaurant meal demand in a new model with machine learning in order to obtain a viable prediction model to avoid overprojection of demand as a result of food waste or subprojection with the consequence of teachers or students without meals.

Key-words: Artificial Neural Networks. Demand Prediction. Machine Learning. Artificial intelligence. Perceptron Multiple layers.

Lista de ilustrações

Figura 1 – DADOS DE DEMANDA DE SAZONALIDADE DIÁRIA, EM FUNÇÃO DO HORÁRIO. Adaptado de (RUAS, 2012).	26
Figura 2 – Relação entre dados, informação e conhecimento. Retirada de (REZENDE, 2003).	27
Figura 3 – Uma trajetória e um conjunto de trajetórias. Adaptado de (GIANNOTTI; PEDRESCHI, 2008).	28
Figura 4 – Exemplos de trajetórias:(1) bruta, (2) e (3) semânticas. Retirada de (??). . .	31
Figura 5 – Exemplo de IB-SMOT com <i>stops</i> encontrados.(ALVARES et al., 2007) . . .	34
Figura 6 – Exemplo de encontro.(??)	35
Figura 7 – Exemplo de convergência.(??)	35

Lista de tabelas

Tabela 1 – Tabela de trajetórias. Fonte:(??)	28
Tabela 2 – Plano de atividades para o TCC II	39

Lista de abreviaturas e siglas

ICT	Instituto Central de Tecnologia
UNIFESP	Universidade Federal de São Paulo

Sumário

1	Introdução	21
1.1	Contextualização e Motivação	21
1.2	Definição do problema	22
1.3	Justificativas	22
1.4	Objetivos	22
1.4.1	Objetivo geral	22
1.4.2	Objetivos específicos	23
1.5	Metodologia	23
1.6	Organização do documento	23
2	Fundamentação Teórica	25
2.1	Introdução	25
2.2	Dados	25
2.3	Séries Temporais	26
2.3.1	Sazonalidade	26
2.3.2	Dados espaço-temporais	27
2.3.3	Informação	27
2.3.4	Conhecimento	27
2.4	Trajectoria	28
2.4.1	Trajectoria de objetos móveis	28
2.4.2	Ponto	29
2.4.3	Objeto	29
2.4.4	Dados semânticos	29
2.4.5	Enriquecimento semântico	29
2.4.6	Semântica em trajetórias	29
2.4.6.1	O modelo <i>stops and moves</i>	30
2.4.6.2	Trajectoria semântica	31
2.5	O algoritmo SMoT	31
2.5.1	Padrões de movimentos em trajetórias	34
3	Trabalhos relacionados	37
4	Plano de atividades para o TCC II	39
5	Conclusão	41
	Referências	43

1 Introdução

1.1 Contextualização e Motivação

A previsão de demanda é um ponto de extrema importância para qualquer empresa, uma previsão adequada permite o ajuste de todo o seu mecanismo de operações para atender tal demanda com a melhor eficiência possível, maximizando lucros, minimizando perdas, e principalmente atendendo todas as necessidades do cliente.

Todo restaurante universitário enfrenta problemas de previsão de demanda de refeições e prejuízos com a falta de vendas e ou o descarte de refeições não vendidas. Um dos grandes problemas enfrentados hoje no mundo é a elevação dos preços dos alimentos, o valor além de monetário é moral, a alimentação é o recurso primitivo de base da humanidade que hoje ainda enfrenta um mal acesso a este em muitas regiões carentes. O descarte indevido de alimentos, provocado por suas limitações e durabilidade não gera somente prejuízos monetários e sim prejuízos ambientais e morais. Isso tem causado preocupações para a população em geral e também para empresas como restaurantes que sofrem diretamente os reflexos da variação no preço dos alimentos e na demanda. Atualmente o restaurante universitário do ICT - UNIFESP não possui um sistema que ajude na gestão de compras dos alimentos.

No restaurante universitário do ICT UNIFESP as refeições são fornecidas de segunda a sexta feira. O caso particular de restaurantes universitários envolve um fluxo de demanda influenciado por dia da semana, visto que a demanda é influenciada pela quantia de alunos presentes na universidade, que por sua vez é influenciada pela grade de aulas determinada semestralmente por dia da semana. O caso de análise para este projeto foi motivado após informações de relevantes desperdícios.

Devido às condições burocráticas no ambiente do restaurante que compreendem fidelidade de contrato, exclusividade na região pois o restaurante se encontra em localização que o faça ser o único provedor de alimentos ao público do campus estando isolado fisicamente de qualquer região comercial, e acessibilidade do público à aquisição de refeições que em sua maior parte adquirem refeições pelo valor de 2,50 sendo o restante do custo da refeição subsidiada pela instituição através de verbas federais, a escolha dos parâmetros não será influenciada por muitos fatores externos como concorrência, acessibilidade do ponto, entre outros.

Outro ponto importante é a obtenção dos valores de venda; não foram escolhidas as vendas diretas do ponto de venda de tickets de refeição, e sim os dados de coleta da entrada do

restaurante, que demonstram a real movimentação de público no restaurante em determinado dia.

O estudo da relação de vendas, temperatura, outras variáveis climáticas e do ambiente, já é comum em outros cenários, entre eles o de maior destaque é na demanda de energia elétrica. Os cenários de vendas de alimentos perecíveis ganha também destaque apesar de se encontrar investimentos maiores na indústria de produção de energia elétrica. O objetivo tanto no cenário deste trabalho, o restaurante universitário, como em outros cenários é o mesmo, atender toda a demanda de consumo e evitar transtorno à qualquer consumidor pela falta da mesma, e evitar prejuízos de produção não consumida. Tais prejuízos impactam não só ao fornecedor, mas também ao consumidor, um fornecimento de produto e serviço com um bom planejamento de demanda, poupa recursos ao produtor que podem ser investidos em melhor qualidade de produto, e menor preço ao consumidor a fim de que seja obtido um modelo de previsão viável para evitar sobrestimação de demanda com consequência de desperdício de alimentos, ou subestimação com consequência de docentes ou discentes sem refeições.

Tal problema tem sido impactante e frequente para o restaurante que informa que em alguns dias no mês passa por sobrestimação e desperdício superior a 100 refeições.

1.2 Definição do problema

O problema a definir neste trabalho é encontrar um modelo de previsão de demanda através de algoritmos de regressão por aprendizado de máquina.

1.3 Justificativas

As atuais abordagens de previsão do restaurante universitário, que envolvem média aritmética simples e dedução subjetiva são falhos por não serem calculados tendo uma visão ampla de todo um histórico de dados de grande amostragem, e também não cruzam informações diversas como dados climáticos, dados de calendário anual, feriados próximos, entre outros.

1.4 Objetivos

1.4.1 Objetivo geral

Construir um modelo utilizando uma Rede Neural Artificial para a previsão da demanda de refeições do restaurante universitário do ICT-UNIFESP com menos de 10

1.4.2 Objetivos específicos

- Construir modelos utilizando três algoritmos de regressão;
- Realizar análise estatística da qualidade dos modelos;

1.5 Metodologia

Por se tratar de um trabalho de previsão de demanda, a primeira atividade será o estudo da tarefa de análise de regressão com foco na previsão de demanda. Análises de regressão fundamentadas por BOCANEIRA (2002) e técnicas aplicadas de regressão propostas por SMOLA (1998), são as principais bases de fundamentação dos trabalhos de previsão de demanda de LOPES(2008) no restaurante universitário da Universidade Federal de Viçosa, e de RUAS et al. (2012) na previsão de demanda de energia elétrica utilizando redes neurais artificiais e regressão por máquinas de vetores de suporte, respectivamente. A partir destes trabalhos, serão estudados diferentes métodos de regressão e como a partir deles podem ser aplicadas técnicas de previsão de demanda por meios estatísticos e usando técnicas de aprendizado de máquina.

Após essa etapa, será então iniciado o estudo aprofundado sobre Redes Neurais Artificiais de acordo com BRAGA (2000), e o estudo específico da técnica de Perceptron Multi-Camadas fundamentado em HAYKIN (1994) que embasa o trabalho realizado na Universidade Federal de Viçosa.

1.6 Organização do documento

Este trabalho está organizado da seguinte forma: o capítulo 2 apresenta os conceitos básicos para o entendimento do trabalho. O capítulo 3 apresenta trabalhos relacionados. O capítulo 4 demonstra o plano de atividades para o TCC II. O capítulo 5 conclui este trabalho.

2 Fundamentação Teórica

2.1 Introdução

A previsão da demanda é o fator principal da eficiência de qualquer modelo de processamento do tipo entrada-saída, onde a sua saída deve atender uma demanda não determinística. É necessário prever a demanda para projetar e aperfeiçoar o processamento e a entrada deste modelo.

O modelo a ser analisado neste trabalho é a linha produtiva de um restaurante universitário, onde o mesmo precisa projetar sua compra de insumos e alocação de recursos na entrada de seu modelo de negócio, e projetar sua saída, que é a produção de refeições em quantidade numérica e inteira distribuída em função do tempo em dias, para atender um consumo humano que não se comporta de maneira determinística, já que este consumo é facultativo ao público do local.

Logo é necessário entender e descobrir quais elementos influenciam este consumo humano, de que forma estes elementos exercem tal influência e em qual intensidade ela ocorre.

E para explorar tais elementos, é necessário explorar como este consumo ocorreu historicamente a fim de se encontrar padrões de comportamento, que no cenário deste trabalho acontece com informações distribuídas através do tempo.

Neste capítulo o principal objetivo é definir todos os conceitos que serão necessários para o entendimento total deste trabalho, analisando a forma que os dados são entendidos, e as ferramentas com os quais possam ser trabalhados.

2.2 Dados

Abstraindo o histórico de vendas do restaurante, temos 2 tipos de dados principais. A venda podendo ser uma observação Y , e a data da venda sendo uma observação X . Se comportando de modo que a função de X implica em Y .

$$X, Y \rightarrow Y \simeq f(X)$$

A observação principal a ser analisada neste trabalho é o numero de vendas do R.U do ICT - UNIFESP, em valor inteiro, obtido de um dia letivo, em um determinado período sendo este o almoço ou janta.

A observação principal tem então 2 tipos, o almoço e a janta, tratados de maneira distinta e não correlacionada neste trabalho, sendo assim, cada um destes 2 tipos, podendo ser explorados com análises e métodos diferentes.

O tempo, onde a informação principal se distribui, é em formato de data, e contém apenas 1 valor da venda total de almoços nesta data, e da venda total de jantas nesta data.

2.3 Séries Temporais

De acordo com (MORETTIN, 1987) uma série temporal é um conjunto de observações ordenadas em função do tempo, comumente iguais. Essas observações apresentam uma dependência serial entre elas, sendo um dos objetivos do estudo de séries temporais, analisar e modelar essa dependência. Além disso, séries temporais são analisadas pelos seus principais movimentos, como tendência, sazonalidade e a componente aleatória, sendo a tendência e sazonalidade as propriedades que mais ganham destaque em pesquisas de previsão de demanda. Os meios mais comuns de se analisar a componente sazonal são o Método de Regressão (paramétrico), Método de Médias Móveis (não paramétrico), e Método de Diferença Sazonal (sazonalidade determinística). Neste trabalho, a distribuição dos dados do Restaurante se dá de forma paramétrica, ou seja, as vendas do restaurante se distribuem em função do tempo e com influências de parâmetros como o dia da semana por exemplo, sendo ideal as análises de regressão.

Figura 1-Demanda em três dias consecutivos

Através da Figura 1 percebe-se a sazonalidade diária, com picos de consumo próximo às 09:00, 14:00 e 20:00 horas, e baixo consumo durante a madrugada, entre 00:00 e 06:00 horas.

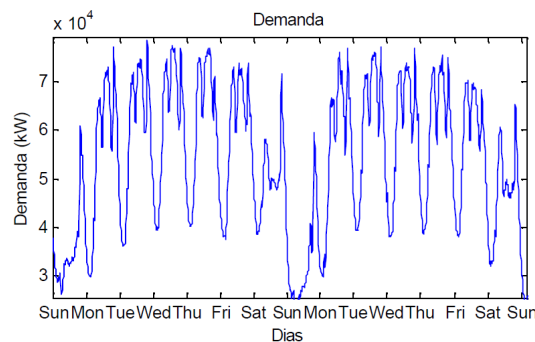


Figura 2-Demanda em duas semanas consecutivas

IMAGEM RF3-01

Figura 1 – DADOS DE DEMANDA DE SAZONALIDADE DIÁRIA, EM FUNÇÃO DO HORÁRIO. Adaptado de (RUAS, 2012).

2.3.1 Sazonalidade

A série temporal sazonal, ocorre com um padrão de mudança que tem uma repetição por ano, podendo ser uma sazonalidade trimestral, semestral, entre outros. No cenário deste estudo a frequência de alunos dentro do ICT - UNIFESP e consequentemente dentro do R.U tem sazonalidade semestral. As técnicas de previsão para evidências sazonais usam métodos de

regressão, que pode ser observados nos trabalhos de previsão de consumo de energia elétrica realizados por (JUNIOR, 2007) ()

2.3.2 Dados espaço-temporais

Dados espaço-temporais chamados de trajetórias brutas, são coletados através de dispositivos móveis. Uma trajetória bruta é representada por um conjunto de pontos (tid, x, y, t) , em que tid é o identificador da trajetória e x, y , são coordenadas geográficas correspondentes a um lugar no espaço em determinado tempo t .

2.3.3 Informação

A informação é o dado analisado e seu contexto ser bem definido compreendendo seus padrões, onde deve-se ter um conjunto de dados a ser processado que podem gerar um valor adicional além dos próprios fatos obtidos.

2.3.4 Conhecimento

Em (REZENDE, 2003), o conhecimento refere-se à habilidade de se criar um modelo mental que descreva o objeto e indique as ações a serem implementadas e as decisões a tomar. A compreensão e análise se tornam necessárias para a tomada de decisões inteligentes, onde são realizadas a partir do nível do conhecimento obtido através das etapas de obtenção dos dados e sua compreensão, retirar informações e obter conhecimento.

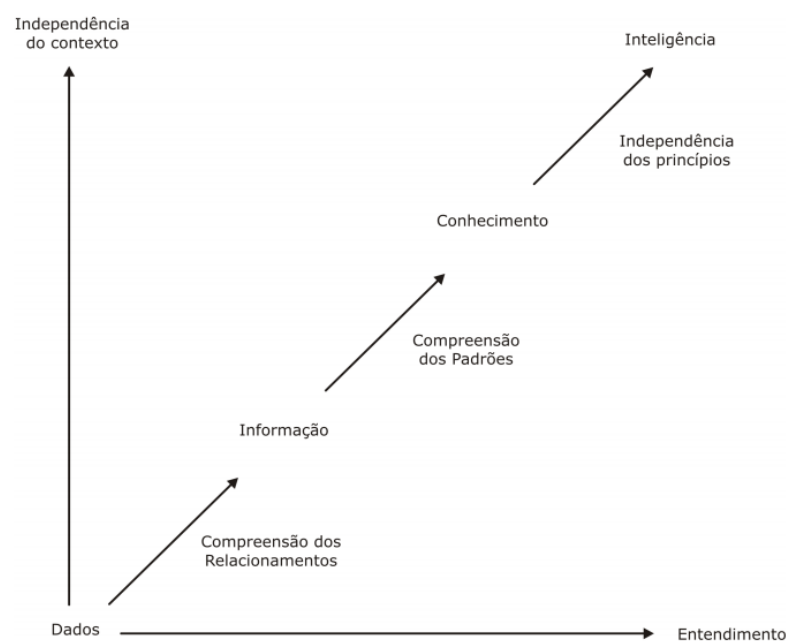


Figura 2 – Relação entre dados, informação e conhecimento. Retirada de (REZENDE, 2003).

2.4 Trajetória

Trajetórias podem ser geradas por qualquer objeto que carregue um dispositivo, seja um pedestre, um carro ou até um animal. Cada trajetória tem um conjunto de propriedades que descrevem o movimento de um objeto. Segundo (GIANNOTTI; PEDRESCHI, 2008), tais propriedades podem ser divididas em características momentâneas e gerais. As momentâneas são o objeto em um determinado tempo t ou determinado momento de um tempo t . As características gerais referem-se a trajetória como um todo, como a forma geométrica, o tamanho total da trajetória, a duração e velocidade média e etc. Podemos definir então trajetória como um segmento espaço-temporal do caminho percorrido por um objeto como definido por (??).

Na figura 2 demonstra exemplos de trajetórias brutas, de apenas um objeto e de vários objetos gerados por um dispositivo móvel, já na tabela 1 mostra como é a forma destes dados gerados.

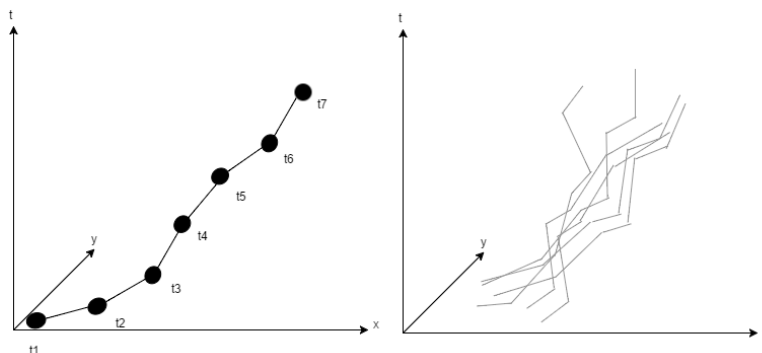


Figura 3 – Uma trajetória e um conjunto de trajetórias. Adaptado de (GIANNOTTI; PEDRESCHI, 2008).

Tabela 1 – Tabela de trajetórias. Fonte:(??)

Tid	Posição(x,y)	Tempo(t)
1	48.890018 2.246100	10:18
1	48.890020 2.246102	10:19
1	48.888880 2.248208	10:20
1	48.885732 2.255031	10:21
1	48.878433 2.283517	10:22
1	48.863214 2.279415	10:23

2.4.1 Trajetória de objetos móveis

Em (??), um aparelho de GPS pode ser configurado para gravar novos dados e consegue, por exemplo, gravar um ponto a cada segundo em que o objeto se move, um ponto a cada uma determinada distância em que o objeto se move ou ainda gravar pontos em determinado tempo t , portanto a sequência dos pontos registrados para cada indivíduo é chamada de trajetória do objeto móvel.

2.4.2 Ponto

É um registro (x, y, t) , no qual x e y são as coordenadas geográficas e t é a identificação de tempo em que a coordenada foi coletada (??).

2.4.3 Objeto

É o objeto em movimento e que transporta um dispositivo móvel. Pode ser uma pessoa, um carro, um animal, um robô etc (??).

2.4.4 Dados semânticos

São dados que buscam complementar as informações referentes às trajetórias, agregando valor, por exemplo, meio de transporte, velocidade, informações climáticas, entre outras (??).

2.4.5 Enriquecimento semântico

Em (FURLETTI et al., 2013) é definido que o processo de agregar informações semânticas às trajetórias brutas é chamado enriquecimento semântico e, gera como saída uma trajetória semântica. Por exemplo, eventos, comportamento, objetivo, ambiente e atividade são formas de enriquecer uma trajetória bruta.

2.4.6 Semântica em trajetórias

Em trajetória bruta onde os dados são gerados por dispositivos móveis como GPS, é difícil extrair conhecimento interessante, entretanto se enriquecermos essas trajetórias com informações que agregam valor aos dados coletados pelos dispositivos de localização então a aplicação de extração de dados poderá ser mais eficiente em seus objetivos. Os benefícios de dados semanticamente enriquecidos podem alcançar resultados onde trajetórias brutas não permitiriam ter essa potencial visibilidade de prever ou tomar alguma decisão, por exemplo, se uma pessoa parou em um local por um determinado período de tempo (e.g. shopping, hotel, supermercado), permite observar qual o deslocamento de certas trajetórias como pessoas que estão hospedadas em um hotel e percorrem a cidade buscando por pontos turísticos e então a partir destes pontos quais restaurantes são mais visitados por essas pessoas.

A idéia de agregar semântica às trajetórias de objetos móveis para facilitar a análise deste tipo de dado foi introduzida pela professora Vânia Bogorny e professor Luís Otávio Alva- res ambos da UFSC em conjunto com o grupo da universidade de Hasselt, um pesquisador da Universidade de Buenos Aires e pesquisadores da EPFL. Foi criado um modelo para adicionar semântica às trajetórias, a fim de facilitar e otimizar consultas espaço-temporais (ALVARES et al., 2007), com o algoritmo SMoT. Os autores teve como base de seu trabalho o modelo

conceitual para trajetórias elaborado pelo grupo de pesquisas do professor Spaccapietra, (??), chamado *stop and moves* como (??) diz em seu livro, uma trajetória bruta pode ser enriquecida com diferentes informações semânticas, de acordo com o contexto da aplicação e o problema que o usuário pretende resolver, ao tentarmos analisar uma trajetória bruta, será muito difícil encontrar algum padrão novo e interessante. Os padrões semânticos podem ser independentes de localização geográfica (se tiverem em lugares esparsos ou com poucos pontos ou trajetórias, em vez de locais densos, com muitos pontos ou trajetórias) e não apresentar nenhuma similaridade. Então padrões semânticos dificilmente serão extraídos de trajetórias brutas, depois de uma análise dos dados não se pode enriquecer semanticamente as trajetórias.

A área de pesquisa que considera contexto ou semântica em trajetórias ainda está na sua infância (??).

2.4.6.1 O modelo *stops and moves*

Até 2006 os estudos de trajetórias consideravam basicamente a geometria das trajetórias. No fim de 2006, publicado oficialmente em 2008, o trabalho de (??), criou um modelo conceitual para trajetórias chamado *stops and moves*.

Este modelo proposto define os *stops* como as partes importantes da trajetória do ponto de vista da aplicação, ou seja, nas paradas de uma trajetória onde a mesma permaneceu por um determinado intervalo de tempo em lugar específico é onde o *stop* tem mais importância, já que certas informações de semânticas podem dizer a qual lugar esta parada indica. Um *move* é definido como a parte da trajetória (ou a subtrajetória) onde liga um *stop* qualquer até o seu próximo *stop* consecutivo, e ainda a parte da trajetória até encontrar um *stop* e a outra parte quando após o último *stop* da trajetória, ou seja, entre o instante inicial e final da trajetória caso não exista nenhum *stop*. Nesse modelo, todas as partes da trajetória que não são *stops* representam *moves*.

Ao considerarmos, por exemplo, um hotel como um *stop*, onde o objeto móvel é uma pessoa que ficou um determinado tempo t neste hotel, no ponto de vista da aplicação o objeto ficou parado no *stop*, pois se mostra relevante a permanência da pessoa no hotel por um determinado tempo.

O processo de *stops and moves* é muito similar à engenharia reversa de banco de dados, neste processo precisamos dos dados para que o modelo *stops and moves* seja gerado a partir desses dados.

(??) define *stops and moves* como a seguir:

- *Stop*:
 - O usuário explicitamente define parte do percurso em início e término para representar um *stop*.
 - O tempo de um percurso de início e término é um intervalo de tempo não nulo.

- O tempo de todas as paradas são separados, ou seja, o tempo de duas paradas são sempre disjuntas.
- *Move*:
 - Uma parte é delimitada por duas extremidades que representam dois *stops* consecutivos, ou o início e o primeiro *stop*, ou o último *stop* e término, ou [início, término] (sem *stops*).
 - O tempo de um percurso de início e término é um intervalo de tempo não nulo.
 - O alcance espacial da trajetória de início e término de um intervalo é o espaço-temporal de uma linha (não um ponto) definido pela função da trajetória.

2.4.6.2 Trajetória semântica

Trajétoria bruta tem suas limitações citadas ao longo de este trabalho, trabalhar com trajetória semântica tem vantagem por ser mais receptivas e compreendidas, onde são fáceis de perceber a diferença entre essas trajetórias, pois semânticas são carregadas com informações ao longo da trajetória, abaixo na figura 3 existem 3 trajetórias (1) Bruta e (2) e (3) semânticas, é fácil diferenciá-las, porém na trajetória 2 se trata de uma aplicação de turismo, pois os pontos de interesse são lugares onde faz parte do contexto da aplicação. Já na trajetória 3 a semântica faz parte de uma aplicação de trânsito e as partes mais interessantes são partes onde se tem o maior movimento de veículos parados. Pontos importantes da trajetória que identificamos na figura 3 onde nas trajetórias (2) e (3) são lugares marcados com cores, significa pontos de *stops*, e o restante dessas trajetórias são *moves*.

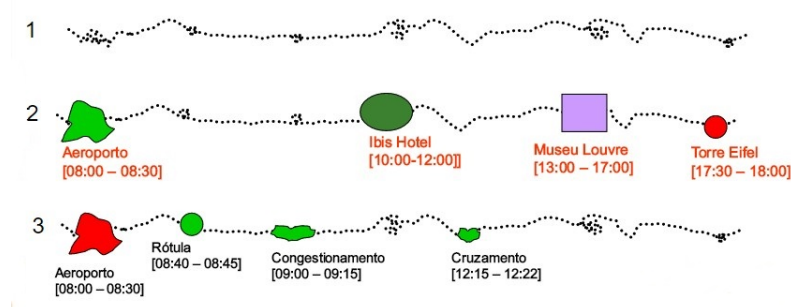


Figura 4 – Exemplos de trajetórias:(1) bruta, (2) e (3) semânticas. Retirada de (??).

2.5 O algoritmo SMoT

A tarefa desempenhada por este algoritmo é verificar se cada ponto de uma trajetória bruta T tem uma intersecção com a área geométrica de um candidato a *stop*, ou seja, essa área geométrica são objetos geográficos relevantes para a aplicação. Se existe a validação de intersecção em algum ponto da trajetória, então o algoritmo verifica a duração da intersecção, que

deve ser igual ou superior que o tempo mínimo estabelecido para aquele tipo de objeto geográfico, então o objeto será considerado um *stop*, em caso de o tempo não ter o seu mínimo atingido aquele objeto geográfico não pode ser considerado um *stop* então não é adicionado aos *stops* encontrados. Este algoritmo possui como entrada um conjunto de trajetórias e um conjunto de candidatos à *stop*, tendo como saída uma tabela de *stops* e uma tabela de *moves*. No Algoritmo 1 é apresentado um pseudocódigo de extração de *stops and moves*, proposto por (ALVARES et al., 2007).

Algoritmo 1: PSEUDOCÓDIGO SMoT**Entrada:** T, A

```

1  início
2       $S \leftarrow newStops()$ 
3       $M \leftarrow newMoves()$ 
4      para cada trajetria  $t \in T$  faça
5           $i \leftarrow 0$ 
6           $previousStop \leftarrow null$ 
7          enquanto  $i \leq size(t)$  faça
8              se  $\exists (Rc, \Delta c) \in A \mid geometria(t[i]) \cap Rc$  então
9                   $enterTime \leftarrow time(t[i])$ 
10                  $i++$ 
11                 enquanto  $geometria(t[i]) \cap Rc$  faça
12                      $i++$ 
13                 fim
14                  $i--$ ,  $leaveTime \leftarrow time(t[i])$ 
15                 se  $leaveTime - enterTime \geq \Delta c$  então
16                      $stop \leftarrow (t, Rc, enterTime, leaveTime)$ 
17                      $S.add(stop)$ 
18                      $move \leftarrow$ 
19                          $(t, previousStop, stop, previousStop.leaveTime, enterTime)$ 
20                      $M.add(move)$ 
21                      $previousStop \leftarrow stop$ 
22                 fim
23                  $i++$ ,  $j \leftarrow 1$ 
24                 enquanto  $(i + j \leq size(t))$  and  $(t[i + j] - t[i] < \min \Delta c(A))$  faça
25                      $j++$ 
26                 fim
27                 se  $\nexists (Rc, \Delta c) \in A \mid geometria(t[i + j - 1]) \cap Rc$  então
28                      $i \leftarrow i + j$ 
29                 fim
30             fim
31             se  $t[i - 1] \text{ not } \in previousStop$  então
32                  $move \leftarrow (t, previousStop, null, previousStop.leaveTime, time(t[i - 1]))$ 
33                  $M.add(move)$ 
34             fim
35         fim
36     fim

```

O algoritmo verifica para cada ponto de uma trajetória t se cruza com a geometria de um candidato a *stop* R_c (linha 8). Em caso afirmativo, o algoritmo avalia a duração da intersecção t com R_c , pelo menos, um limite igual a Δc (linha 15). Se esse é o caso esse *stop* é gravado. Note que quando usamos t e R_c como parâmetros (por exemplo, linha 16), usamos e atribuímos na realidade um identificador de chave estrangeira e não toda uma estrutura de dados.

Esse método de *stops and moves* tem um nome mais específico chamado IB-SMoT (*Intersection-Based Stops and Moves of Trajectories*) que por meio de sobreposição dos principais pontos de uma trajetória bruta, respeitando o tempo limite mínimo de intersecção do ponto, se obtêm a partir de locais preestabelecidos pelo usuário candidatos de *stops* na trajetória. Na figura 4 o algoritmo IB-SMoT encontrou três candidatos a *stop* na trajetória.

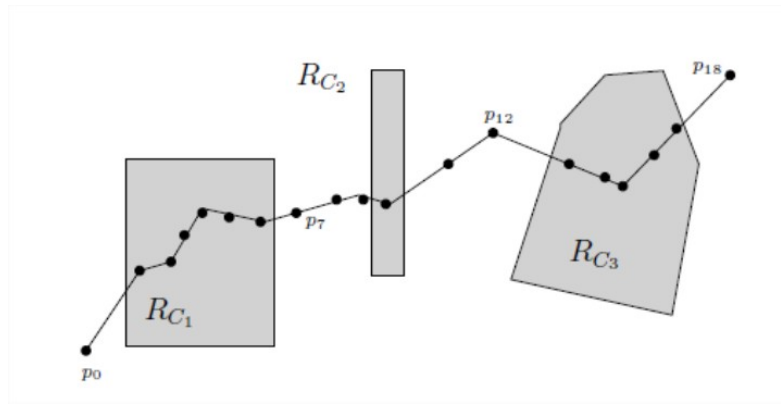


Figura 5 – Exemplo de IB-SMOT com *stops* encontrados. (ALVARES et al., 2007)

2.5.1 Padrões de movimentos em trajetórias

Um movimento relativo de várias trajetórias de objetos móveis pode auxiliar a encontrar certos padrões em trajetórias como recorrências e encontros que é o principal foco deste trabalho, em (LAUBE; KREVELD; IMFELD, 2005) descreve esses padrões de movimentos como a seguir:

- Encontro é um determinado número de trajetórias dentro de um raio R , onde um conjunto m de trajetórias em um determinado intervalo i se move em uma mesma direção, passando por uma determinada área fazendo intersecção com o intervalo de raio R tendo uma mesma velocidade.
- Convergência é um conjunto m de trajetórias que interceptam uma mesma região de raio R , não necessariamente no mesmo momento.

Observe a diferença entre estes dois padrões de movimento, na figura 5 o padrão é de encontro, onde as trajetórias têm a mesma direção e velocidade e ficam dentro do raio um certo tempo. Já na figura 6 o padrão é de convergência, onde as trajetórias vêm de todos os sentidos na mesma direção, porém não ao mesmo tempo.

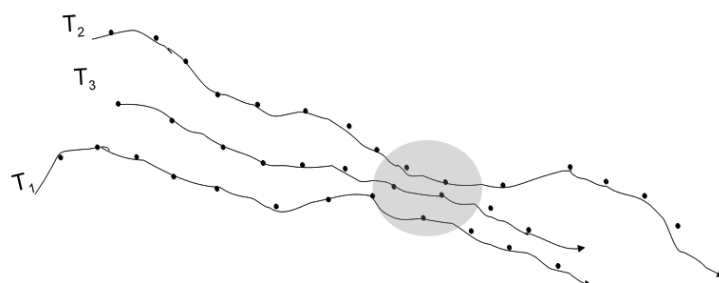


Figura 6 – Exemplo de encontro.(?)

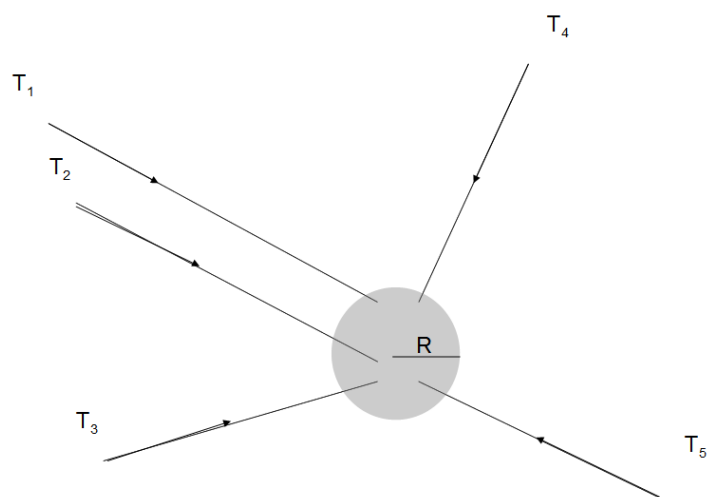


Figura 7 – Exemplo de convergência.(?)

3 Trabalhos relacionados

Entre os últimos artigos publicados em 2015 sobre enriquecimento de trajetórias está o trabalho de (SUBLIME et al., 2015), que tem características específicas que permitem a extração de informações semântica sobre dados geográficos de imagens de alta resolução, esse tipo de extração de semântica é muito usado quando os dados não são obtidos via objetos móveis, já que se obtêm a semântica a partir de dados. Porém com estes dados, são gerados mapas atualizados que são proveitosos para aplicar semântica em trajetórias brutas de objetos móveis. Os estudos em torno de trajetórias semânticas e enriquecimento semântico têm tido um grande avanço nos últimos anos.

No trabalho de (ALVARES et al., 2007), o enriquecimento de trajetórias é feito com informações geográficas para simplificar consultas que são complexas, então é proposto um pré-processamento de modelo de dados para adicionar informações semânticas as trajetórias.

Em (LAUBE; KREVELD; IMFELD, 2005), é introduzido o estudo de comportamento de movimentos de trajetórias, onde o autor explica como se obtêm cada tipo de movimento o que vem a ser proveitoso para este trabalho, o padrão de encontro que é exemplificado no livro é o mais complexo de se calcular, por que temos que medir o intervalo de tempo de todas as trajetórias que passam em um determinado raio R .

Em (??) uma trajetória bruta pode ser enriquecida com diferentes informações semânticas de acordo com o contexto da aplicação e o problema que o usuário pretende resolver.

Neste trabalho é utilizado o método de enriquecimento de trajetórias onde usaremos informações geográficas para conflitar com as trajetórias e dizer se são pontos de *stops* através do algoritmo de *stops and moves*, tendo esses dados enriquecidos semanticamente vai ser implementado um algoritmo que detecta encontros em trajetórias a partir de um raio dado e um tempo de permanência mínimo em cada *stop*.

4 Plano de atividades para o TCC II

Para a continuação deste trabalho o cronograma abaixo deve ser seguido.

Tabela 2 – Plano de atividades para o TCC II

Atividades	Março	Abril	Maio	Junho	Julho
1	✓				
2	✓				
3		✓	✓		
4		✓	✓	✓	
5		✓	✓	✓	
6					✓

1. Busca de uma base de dados correta para a aplicação.
2. Evidenciar o contexto semântico a ser aplicado.
3. Implementar o algoritmo SMoT.
4. Implementar o algoritmo que detecta encontros.
5. Análise e comparação dos resultados.
6. Documentação total do trabalho.
7. Apresentação do trabalho.

5 Conclusão

A forma de enriquecer semânticas de objetos móveis, é uma área de pesquisa que cresce cada vez mais, e a produção de dispositivos móveis cresce de uma forma muito agressiva no mercado resultando em grandes volumes de dados geográficos, porém aonde aplicar todos os métodos que vimos neste trabalho e obter resultados significativos, isso ainda está em sua infância.

Em aplicações onde se usam dados gerados por dispositivos móveis, deve ser feito um estudo muito bem-feito para enriquecer suas trajetórias, onde o contexto da aplicação depende muito na execução do algoritmo para obter excelentes resultados.

A proposta deste trabalho é trabalhar com dados fornecidos de trajetórias brutas e aplicar o método SMOt para enriquecê-las semanticamente e posteriormente analisar essas rotas e indicar em quais pontos houve encontros, independente de qual objeto gerou tais rotas.

Referências

- ALVARES, L. O. et al. A model for enriching trajectories with semantic geographical information. In: *Proceedings of the 15th Annual ACM International Symposium on Advances in Geographic Information Systems*. New York, NY, USA: ACM, 2007. (GIS '07), p. 22:1–22:8. ISBN 978-1-59593-914-2. Disponível em: <<http://doi.acm.org/10.1145/1341012.1341041>>. Citado 5 vezes nas páginas 13, 29, 32, 34 e 37.
- FURLETTI, B. et al. Inferring human activities from gps tracks. In: *Proceedings of the 2Nd ACM SIGKDD International Workshop on Urban Computing*. New York, NY, USA: ACM, 2013. (UrbComp '13), p. 5:1–5:8. ISBN 978-1-4503-2331-4. Disponível em: <<http://doi.acm.org/10.1145/2505821.2505830>>. Citado na página 29.
- GIANNOTTI, F.; PEDRESCHI, D. *Mobility, data mining and privacy: Geographic knowledge discovery*. [S.l.]: Springer Science & Business Media, 2008. Citado 2 vezes nas páginas 13 e 28.
- JUNIOR, A. M. *Análise de métodos de previsão de demanda baseados em séries temporais em uma empresa do setor de perfumes e cosméticos*. Dissertação (Mestrado) — Pontifícia Universidade do Paraná. Curitiba, 2007. Citado na página 27.
- LAUBE, P.; KREVELD, M. van; IMFELD, S. Finding remo—detecting relative motion patterns in geospatial lifelines. In: *Developments in spatial data handling*. [S.l.]: Springer Berlin Heidelberg, 2005. p. 201–215. Citado 2 vezes nas páginas 34 e 37.
- MORETTIN, P. A. *Análise de séries temporais*. [S.l.]: Editora Blucher, 1987. Citado na página 26.
- REZENDE, S. O. *Sistemas inteligentes: fundamentos e aplicações*. [S.l.]: Editora Manole Ltda, 2003. Citado 2 vezes nas páginas 13 e 27.
- RUAS, G. I. e. S. Utilização de redes neurais artificiais para a determinação do números de refeições diárias de um restaurante universitário. Universidade Federal do Paraná(UFPR), n. 1, p. 10, 2012. Disponível em: <http://www.dcc.fc.up.pt/~ines/enia07_html/pdf/27927.pdf>. Citado 2 vezes nas páginas 13 e 26.
- SUBLIME, J. et al. Semantic rich icm algorithm for vhr satellite images segmentation. In: *Machine Vision Applications (MVA), 2015 14th IAPR International Conference on*. [S.l.: s.n.], 2015. p. 45–48. Citado na página 37.