

IV Workshop de Dissertações de Mestrado do PPgSI (2017)

FICHA DA PESQUISAⁱ

DADOS GERAIS				
Título do projeto de pesquisa	Predição de tempo de chegada de ônibus usando colaboração coletiva			
Orientando	Diego Vieira Neves			
Orientador (es)	Dr. Daniel de Angelis Cordeiro			
Semestre no curso, na data do workshop	<input type="checkbox"/> 2º semestre	<input type="checkbox"/> 3º semestre	<input checked="" type="checkbox"/> 4º semestre	<input type="checkbox"/> 5º semestre
Qualificação	<input checked="" type="checkbox"/> Qualificação já realizada em: 28/06/2017 <input type="checkbox"/> Realização da qualificação planejada para: dd/mm/aaaa			
Defesa	Prazo máximo para depósito: 29/08/2018 Realização da defesa planejada para: dd/mm/aaaa			
Linha e Área de pesquisa	Gestão e desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Gestão de SI <input checked="" type="checkbox"/> Eng. de Software <input type="checkbox"/> IHC		Inteligência de Sistemas: <input type="checkbox"/> IA <input type="checkbox"/> Rec. de Padrões <input type="checkbox"/> Proc. Gráfico	
Área de aplicação	<input type="checkbox"/> Ambientes Corporativos	<input type="checkbox"/> Educação	<input type="checkbox"/> Linguagem Natural	<input type="checkbox"/> Redes Sociais
	<input type="checkbox"/> Bioinformática	<input type="checkbox"/> Educação a Distância	<input type="checkbox"/> Linguística	<input type="checkbox"/> Robótica
	<input type="checkbox"/> Biometria	<input type="checkbox"/> Internet	<input type="checkbox"/> Processos de Negócio	<input type="checkbox"/> Saúde
	<input type="checkbox"/> Economia	<input type="checkbox"/> Jogos	<input type="checkbox"/> Química	<input checked="" type="checkbox"/> Smart Cities
Publicações associadas ao projeto de mestrado	Informe se há publicações associadas a seu projeto de mestrado, deixando claro seu estado: “em elaboração”, “submetida”, “aceita”, “publicada”. Informe veículo e tipo de publicação.			

DESCRIÇÃO DO PROJETO DE PESQUISA	
Contextualização / motivação	Este projeto investiga como o uso de dados obtidos por colaboração coletiva (<i>crowdsourcing</i>) podem aumentar a confiabilidade das informações sobre o transporte coletivo disponibilizados ao usuário do município de São Paulo.
Problema de pesquisa	Esta pesquisa pretende analisar a viabilidade da utilização dos dados obtidos por meio de técnicas de <i>crowdsourcing</i> no enriquecimento dos dados sobre o transporte coletivo fornecidos pelas aplicações de SIU.
Objetivo geral da pesquisa	Esse trabalho tem como objetivo desenvolver uma metodologia capaz de melhorar a predição de tempo de chegada de ônibus com o uso de informações coletadas mediante a utilização de técnicas de <i>crowdsourcing</i> .
Trabalhos relacionados	Os estudos desenvolvidos por Lin, Zeng 1999, Weigang et al. 2002, Chien, Kuchipudi 2003, Chung, Shalaby 2007, Sun et al. 2007 descrevem resultados positivos durante a execução das atividades de predição dos tempos de chegadas de transportes coletivos. Contudo esses trabalhos não consideram variáveis aleatórias tais como, os atrasos provocados em decorrência de congestionamento, número de passageiros, pontos de paradas, etc.
Justificativa e relevância	Um dos problemas mais conhecidos dos transportes coletivos é o do cumprimento de horários. A predição dos tempos de chegadas é uma atividade que depende de vários fatores aleatórios (número de passageiros em pontos de parada, etc.), e desta forma, a sua precisão pode ser prejudicada, pela falta dessas informações. Neste sentido, as técnicas de <i>crowdsourcing</i> possibilita a obtenção de novos dados ou até mesmo a correção de informações existentes.
Proposta para Solução	Está sendo desenvolvida uma métrica para possibilitar a manipulação e a identificação automaticamente de registros que por algum motivo apresente informações erradas. Objetiva-se ainda realizar, a implementações de técnicas que possibilite melhorar a qualidade das atividades de predição, com base em informações obtidas por meio de técnicas de <i>crowdsourcing</i> .
Dados	Os dados utilizados até o momento foram obtidos por uma parceria com a startup Scipopulis, e combinados com os dados que são disponibilizados publicamente pela São Paulo Transporte S. A. (SPTrans). As informações obtidas estão organizadas em duas bases de dados: (i) base de dados SPTrans e (ii) base de dados <i>crowdsourcing</i> .
Validação	A validação será feita por meio de uma análise sobre os resultados obtidos, a fim de verificar se as hipóteses foram de fato comprovadas. Além disso, outro aspecto desta avaliação será a comparação de soluções propostas por outros autores.
Limitações	Um conjunto de dados com informações insuficiente para realização das atividades de predição de tempo de chegada de ônibus.
Resultados esperados	Contribuições científicas: A metodologia que será desenvolvida, poderá auxiliar os pesquisadores da área de análise de dados de mobilidade urbana, dentro do contexto de predição de horários de chegada de ônibus. Contribuições tecnológicas: espera-se que a metodologia desenvolvida ajude a melhorar a confiabilidade das predições de tempo de chegada, contribuindo significativamente no processo de melhoria da qualidade dos serviços disponibilizados para a população.

MÉTODO DE PESQUISA [para cada item, marque todas as opções adequadas para seu projeto]

(basear-se em: <http://www.ufrgs.br/cursopgdr/downloadsSerie/derad005.pdf>)

Gênero	<input type="checkbox"/> Pesquisa teórica	<input checked="" type="checkbox"/> Pesquisa prática	<input type="checkbox"/> Pesquisa empírica	<input type="checkbox"/> Pesquisa metodológica
Natureza	<input type="checkbox"/> Pesquisa básica/pura	<input checked="" type="checkbox"/> Pesquisa aplicada		
Objetivo	<input checked="" type="checkbox"/> Pesquisa exploratória	<input type="checkbox"/> Pesquisa descritiva	<input type="checkbox"/> Pesquisa explicativa	<input type="checkbox"/> [outro – escrever]
Abordagem	<input type="checkbox"/> Pesquisa quantitativa	<input type="checkbox"/> Pesquisa qualitativa	<input checked="" type="checkbox"/> Pesquisa mista (quali-quant)	
Procedimento(s) técnico(s)	<input checked="" type="checkbox"/> Pesquisa experimental	<input type="checkbox"/> Pesquisa com <i>survey</i>	<input type="checkbox"/> Pesquisa netnográfica	
	<input checked="" type="checkbox"/> Pesquisa bibliográfica	<input type="checkbox"/> Estudo de caso	<input type="checkbox"/> Teoria fundamentada em dados (<i>grounded theory</i>)	
	<input type="checkbox"/> Pesquisa documental	<input type="checkbox"/> Pesquisa participante	<input type="checkbox"/> Ciência do projeto (<i>Design science research</i>)	
	<input type="checkbox"/> Pesquisa <i>ex-post-facto</i>	<input type="checkbox"/> Pesquisa-ação	<input type="checkbox"/> [outro – escrever]	
	<input type="checkbox"/> Pesquisa de levantamento	<input type="checkbox"/> Pesquisa etnográfica		
Fonte(s) de dados	<input checked="" type="checkbox"/> pesquisa de laboratório	<input type="checkbox"/> pesquisa de campo	<input checked="" type="checkbox"/> pesquisa bibliográfica	<input type="checkbox"/> [outro – escrever]
Técnica(s) / Instrumento(s) de coleta de dados	<input checked="" type="checkbox"/> medição		<input checked="" type="checkbox"/> observação (direta / participante)	
	<input type="checkbox"/> questionário	<input type="checkbox"/> formulário	<input type="checkbox"/> diário de campo / notas de campo	
	<input type="checkbox"/> entrevista	<input type="checkbox"/> <i>benchmark</i>	<input type="checkbox"/> análise documental (ou de artefatos)	
	<input type="checkbox"/> grupos focais		<input type="checkbox"/> [outro – escrever]	
Técnica(s) de análise de dados	Análise quantitativa: <input checked="" type="checkbox"/> Estatística descritiva <input type="checkbox"/> Estatística inferencial <input type="checkbox"/> [outro – escrever]		Análise qualitativa: <input checked="" type="checkbox"/> Análise de conteúdo <input type="checkbox"/> Análise do discurso <input type="checkbox"/> [outro – escrever]	

CRONOGRAMA [altere conforme necessário]

	2017											2018												
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Estudo da literatura	X	X	X	X																				
Análise da base de dados utilizada		X	X	X	X																			
Levantamento das técnicas mais utilizadas para predição			X	X	X																			
Redação do texto para a qualificação	X	X	X	X	X																			
Exame de Qualificação						X																		
Definição de métricas para identificar dados inconsistentes					X	X	X	X																
Implementações de técnicas						X	X	X	X	X														
Avaliação dos resultados							X	X	X	X	X													
Redação de artigos científicos									X	X	X	X	X											
Redação de texto para a dissertação									X	X	X	X												
Depósito da dissertação													X											
Defesa da dissertação														X										