PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS

Programa de Pós-Graduação em Informática

Faber Henrique Zacarias Xavier

CARACTERIZAÇÃO DA CARGA DE TRABALHO DE UMA REDE DE TELEFONIA MÓVEL EM GRANDES EVENTOS

Belo Horizonte

2014

Faber Henrique Zacarias Xavier

CARACTERIZAÇÃO DA CARGA DE TRABALHO DE UMA REDE DE TELEFONIA MÓVEL EM GRANDES EVENTOS

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Pontifícia Universidade Católica de Minas Gerais, como requisito parcial para obtenção do título de Mestre em Informática.

Orientador: Prof. Humberto T. Marques Neto

Belo Horizonte

2014

FICHA CATALOGRÁFICA
Elaborada pela Biblioteca da Pontifícia Universidade Católica de Minas Gerais

Xavier, Faber Henrique Zacarias
X3c Caracterização da carga de trabalho de uma rede de telefonia móvel em grandes eventos / Faber Henrique Zacarias Xavier. Belo Horizonte, 2014.
94 f. : il.

Orientador: Humberto T. Marques Neto
Dissertação (Mestrado) - Pontifícia Universidade Católica de Minas Gerais. Programa de Pós-Graduação em Informática.

1. Telefonia celular. 2. Carga de Trabalho. 3. Sistemas de comunicação sem fio. 4. Promoção de eventos especiais. I. Marques Neto, Humberto T.. II. Pontifícia Universidade Católica de Minas Gerais. Programa de Pós-Graduação em Informática. III. Título.

SIB PUC MINAS

CDU: 621.395.74

Faber Henrique Zacarias Xavier

CARACTERIZAÇÃO DA CARGA DE TRABALHO DE UMA REDE DE TELEFONIA MÓVEL EM GRANDES EVENTOS

Dissertação apresentada ao Programa de Pós-Graduação em Informática como requisito parcial para qualificação ao Grau de Mestre em Informática pela Pontifícia Universidade Católica de Minas Gerais.

Prof. Humberto Torres Marques Neto –
PUC Minas

Prof.ª Fátima Duarte Figueiredo –
PUC Minas

Prof.ª Cristina Duarte Murta –
CEFET-MG

Belo Horizonte, 11 de novembro de 2014.

AGRADECIMENTOS

William Shakespeare disse: “A gratidão é o tesouro dos humildes”. A minha eterna gratidão é tudo que tenho a oferecer àqueles aqui listados. Inicialmente gostaria de agradecer aos meus pais pelo apoio e suporte durante os anos de desenvolvimento deste trabalho. A minha noiva sempre parceira e disponível a atender os pedidos de revisão, agora entendedora de eventos de larga escala. Aos meus companheiros de caminhada Agripino e Cíntia que batalharam comigo em PAA. Sozinho não seria possível concretizar esse trabalho, sem o trabalho em equipe de meu irmão Wender e de meu amigo Lucas o *PrediTraf* não estaria completo. Também fazem parte dessa equipe os professores Artur Ziviani e Jussara Almeida, o ganho de conhecimento pela presença de vocês foi fundamental para a qualidade do trabalho. Gostaria de deixar meu mais sincero agradecimento ao professor e orientador Humberto Torres, pela confiança em meu trabalho e ter auxiliado na minha formação tanto profissional, pessoal e acadêmica. A busca pela excelência sempre tentando tirar o melhor trabalho que posso desenvolver será algo que levarei para o resto da vida. Agradeço também a toda equipe da Oi Telecom que contribui com este trabalho, em especial na pessoa do Carlos Malab, presente em todas as etapas do desenvolvimento da ferramenta e sempre incentivando a sua evolução. Agradeço a Rosa que sempre resolveu meus problemas burocráticos. Finalmente, agradeço a PUC Minas pela oportunidade e ao professor Henrique Cota sempre incentivador a carreira acadêmica desde os tempos da graduação.

"Feliz aquele que transfere o que sabe e aprende o que ensina."

Cora Coralina

RESUMO

O planejamento eficiente da infraestrutura necessária para eventos de grande porte é uma tarefa desafiadora para as operadoras de telefonia celular. A diversidade de eventos, bem como as características que os seus participantes (usuários que estão na região do evento) possuem, tornam esta tarefa de alta complexidade. Este trabalho apresenta uma metodologia que suporta a caracterização dos participantes desses eventos de larga escala, permitindo assim uma identificação dos padrões de uso dos recursos da rede de telefonia móvel. Neste trabalho são analisados eventos de natureza distintas como partidas da Copa do Mundo de 2014, Festas de Réveillon, Shows e Corridas de Automobilismo. Este trabalho utiliza de métricas como número de chamadas, intervalos entre chamadas e duração para definir características dos usuários durante estes eventos de larga escala. Além do trabalho de caracterização esta dissertação também apresenta estudos sobre a mobilidade dos participantes a partir de uma metodologia específica para tal análise. A utilização dessas métricas permitiu identificar períodos de concentração do uso dos recursos da rede de telefonia móvel. Os resultados também permitiram específicar qual distribuição de probalidade melhor representa como os usuários realizam chamadas e sua duração. Por fim, ao longo do desenvolvimento deste trabalho foi criada uma ferramenta própria para análise da mobilidade humana em grandes eventos. Espera-se que os resultados apresentados possam melhorar a compreensão das características dessa mobilidade humana em decorrência de grandes eventos.

Palavras-chave: Caracterização de carga de trabalho. Mobilidade Humana. Eventos de larga escala

ABSTRACT

Planning the infrastructure for large scale events is a challenging task for mobile phone carriers. The diversity of events, as well as the characteristics of their participants (users that are within the event area), contributes with the complexity of the task. This work presents a methodology that supports the characterization of participants on these large scale events, for allowing the usage patterns identification of mobile network resources due to different events like 2014 World Cup matches, new year's eve parties, concerts and car races. This work uses the following metrics: calls number, interval between calls, and its duration to define cellphone user's behavior during these large scale events. These metrics allowed to identify periods of concentration of resource usage of a telephone network mobile. The results also allowed to specify which probability distribution represents how users make calls and their respective duration. Besides the characterization work, this dissertation shows a study over the mobility of large scale events based on a specific methodology. Finally, during the development of this work we created a tool to analyze the carrier's network workload. We hope that these results could help for understanding the characteristics of human mobility due to large scale events.

Keywords: Workload Characterization. Human Mobility. Large-scale events

LISTA DE FIGURAS

FIGURA 1 – Metodologia Proposta 33
FIGURA 2 – Linha do tempo para a análise da carga durante um evento 35
FIGURA 3 – Cenários de cálculo do T_{real} 37
FIGURA 4 – Linha do tempo para jogos de Futebol 40
FIGURA 5 – Linha do tempo para Shows e Réveillon 40
FIGURA 6 – Linha do tempo para Corridas Automobilísticas 41
FIGURA 7 – Ranking de ERBs mais utilizadas 43
FIGURA 8 – P(x) número de chamadas - Cruzeiro x Atlético 45
FIGURA 9 – Gráficos comparativos - Número de chamadas 47
FIGURA 10 – Efeito “W” 48
FIGURA 11 – Alterações no efeito “W” 49
FIGURA 12 – Número de chamadas - Jogos decisivos no Maracanã 49
FIGURA 13 – Efeito ”W” nos Jogos da Copa de 2014 50
FIGURA 14 – Número de chamadas - Show Paul McCartney 51
FIGURA 15 – Número de chamadas shows 51
FIGURA 16 – Número de chamadas Réveillons 53
FIGURA 17 – Número de chamadas Réveillon Recife 53
FIGURA 18 – Número de chamadas GP do Brasil e Final Copa do Brasil 54
FIGURA 19 – IAT e IDT Finais Futebol 56
FIGURA 20 – Comparação do IAT dos jogos da Copa com o Jogo entre Flamengo e Vasco (28) 56
FIGURA 21 – IAT e IDT festas de Réveillon 57
FIGURA 22 – Comparativo Número de chamadas e IAT do Réveillon de Salvador . 58
FIGURA 23 – Comparativo IAT Automobilismo, Réveillon e Partida de Futebol . . 59

FIGURA 24 – Gráficos comparativos - Duração 61
FIGURA 25 – Duração Chamadas - Finais Maracanã e Flamengo x Vasco (04) 62
FIGURA 26 – Duração jogos de futebol 62
FIGURA 27 – Duração Réveillon - Belo Horizonte Recife e Rio de Janeiro 63
FIGURA 28 – Comparação Réveillon São Paulo e Flamengo x Vasco (28) 64
FIGURA 29 – Duração Réveillon e show 64
FIGURA 30 – Análise Comparativa Número de Chamadas - Engenhão 66
FIGURA 31 – Análise Comparativa Número de Chamadas - Réveillon de Salvador 67
FIGURA 32 – Análise Comparativa Número de Chamadas - GP do Brasil 68
FIGURA 33 – Comparativo Número de chamadas Engenhão Sem evento 68
FIGURA 34 – Comparativo IAT e IDT região de Interlagos 70
FIGURA 35 – Duração Réveillon e show 71
FIGURA 36 – Mobilidade participantes durante o t_{evento} no Mineirão 73
FIGURA 37 – Mobilidade $t_{partida}$ Cruzeiro x Atlético (Verde) - Brasil x Chile (Roxo) - Show Paul McCartney (Azul) 74
FIGURA 38 – Modelo da ferramenta 79
FIGURA 39 – Diagrama de pacotes da ferramenta 79
FIGURA 40 – Frame do vídeo de mobilidade 81
FIGURA 41 – Tela inicial do PrediTraf 83
FIGURA 42 – Tela de resultados do PrediTraf 84
FIGURA 43 – Mapa de calor da mobilidade dos usuários 85

LISTA DE TABELAS

TABELA 1 – Informações das partidas de futebol analisadas 27

TABELA 2 – Informações sobre os shows analisados 29

TABELA 3 – Informações sobre os Réveillons analisados 30

TABELA 4 – Corridas de Automóveis coletadas 31

TABELA 5 – Linha de tempo por evento 41

TABELA 6 – Valores totais das bases 42

TABELA 7 – *P-Value* - Número de Chamadas 45

TABELA 8 – Jogos de futebol – Número de chamadas agregado em intervalos de 15 minutos 46

TABELA 9 – Shows – Número de chamadas agregado em intervalos de 15 minutos 50

TABELA 10 – Réveillon – Número de chamadas agregado em intervalos de 15 52

TABELA 11 – Tabela Automobilismo - Análise agregada 54

TABELA 12 – Tabela IAT e IDT - Jogos de futebol 55

TABELA 13 – Tabela IAT e IDT - Réveillons e show 57

TABELA 14 – Tabela IAT e IDT - Automobilismo 58

TABELA 15 – *P-Value* - Duração das Chamadas 59

TABELA 16 – Tabela de duração - Jogos de futebol 60

TABELA 17 – Tabela de duração - Réveillons e show 63

TABELA 18 – Tabela de duração - Automobilismo 64

TABELA 19 – Valores totais das bases - Partidas de Futebol 65

TABELA 20 – Valores totais das bases de dados- Réveillon 66

TABELA 21 – Valores totais das bases - Automobilismo 67

TABELA 22 – Tabela IAT e IDT - Jogos de futebol 68

TABELA 23 – Tabela IAT e IDT - Réveillon 69

TABELA 24 – Tabela IAT e IDT - Automobilismo 69
TABELA 25 – Tabela Comparativa duração - Engenhão Sem Evento 70
TABELA 26 – Tabela Comparativa duração - Réveillon e Regiões Sem Evento 71
TABELA 27 – Tabela Comparativa duração - Região de Interlagos 71
TABELA 28 – Tabela característica do Modelos do *PrediTraf*. 76
TABELA 29 – Coletas realizadas durante o trabalho 94

SUMÁRIO

1 INTRODUÇÃO **15**

2 REVISÃO BIBLIOGRÁFICA **19**

2.1 Caracterização de Carga de Trabalho 19

2.2 Mobilidade Humana 20

2.3 Uso de Dados de Telefonia para Construção de Cidades Inteligentes 23

3 BASES DE DADOS **25**

3.1 Estrutura da Base de Dados 25

3.2 Eventos Coletados 26

3.2.1 Partidas de Futebol 26

3.2.2 Shows 29

3.2.3 Réveillon 30

3.2.4 Automobilismo 31

4 METODOLOGIA **33**

4.1 Metodologia de Caracterização 33

4.1.1 Tratamento de Dados 34

4.2 Metodologia de Mobilidade 35

5 RESULTADOS **39**

5.1 Aplicação da Metodologia nas Bases de Dados 39

5.1.1 Visão Geral das Bases de Dados 41

5.2 Análise das Características do Evento 43

5.2.1 Número de Chamadas 44

5.2.1.1 Partidas de Futebol 46

5.2.1.2 Shows 50

5.2.1.3 Réveillon 51

5.2.1.4 Automobilismo 54

5.2.2 Intervalo entre Chamadas 54

5.2.2.1 Partidas de Futebol 55

5.2.2.2 Réveillon e Shows 56
5.2.2.3 Automobilismo 58
5.2.3 Duração das Chamadas 59
5.2.3.1 Partidas de Futebol 60
5.2.3.2 Réveillon e Shows 62
5.2.3.3 Automobilismo 63
5.3 Análise Comparativa da Região do Evento 64
5.3.1 Número de Chamadas 65
5.3.1.1 Partidas de Futebol 65
5.3.1.2 Réveillon 65
5.3.1.3 Automobilismo 67
5.3.2 Intervalo entre Chamadas 67
5.3.2.1 Partidas de Futebol 67
5.3.2.2 Réveillon 68
5.3.2.3 Automobilismo 69
5.3.3 Duração 70
5.3.3.1 Partidas de Futebol 70
5.3.3.2 Réveillon 70
5.3.3.3 Automobilismo 71
5.4 Análise da Mobilidade 72

6 PREDITRAF 75
6.1 O Modelo do PrediTraf 76
6.1.1 Modelo de Simulação do PrediTraf 76
6.1.2 Entradas do Modelo 76
6.1.3 Saídas do Modelo 78
6.2 Arquitetura do PrediTraf 78
6.2.1 Módulo de Geração de Mapas e Vídeos 80
6.3 Interfaces 82
6.3.1 Interface Principal 82
6.3.2 Interface de Resultados 83

7 CONCLUSÃO 87

REFERÊNCIAS **89**

8 LISTA DE EVENTOS COLETADOS **94**

1 INTRODUÇÃO

O crescimento das bases de dados da rede telefonia móvel é um fenômeno que vem acontecendo nos últimos anos no Brasil. Segundo dados da Agência Nacional de Telecomunicações (ANATEL), o país atingiu em outubro de 2010 a marca de um celular por habitante, com aproximadamente 195 milhões de telefones móveis. Desde então esses valores vêm aumentando constantemente. Em um relatório emitido em abril de 2014 pela ANATEL, o país possuía 274 milhões de telefones móveis, ou seja, um crescimento de 3,4% em relação ao último ano.

Este crescimento no número de dispositivos móveis, aliado ao aumento da renda da população e a criação de planos populares de telefonia, colaborou para um crescimento na utilização dos serviços de telefonia móvel tanto em volume quanto em tempo de uso. As operadoras têm investido constantemente em melhorias de sua infraestrutura e na qualidade do serviço prestado. Entretanto, um relatório emitido pela Fundação de Proteção e Defesa do Consumidor de São Paulo (Procon-SP) mostrou que o setor de telecomunicações liderou o *ranking* de reclamações no ano de 2013.

A análise dos dados da rede de telefonia móvel gerados pela utilização dos serviços de telefonia móvel podem gerar benefícios não apenas para as operadoras, mas para toda sociedade, como por exemplo, nas áreas de saúde, mobilidade e segurança (GONZALEZ; HIDALGO; BARABASI, 2008; SOPER, 2012). Os resultados dessas pesquisas podem apoiar tomadas de decisões de empresas e entidades governamentais, permitindo a esses grupos agirem com antecedência nas mais diversas situações.

Contudo, as operadoras normalmente encontram dificuldades para planejar o uso de sua rede de telefonia celular em situações que geram uma carga de trabalho fora do padrão, como por exemplo, durante a realização de eventos de larga escala (BLEICHER, 2012). Como exemplos desses grandes eventos, podemos mencionar a partida final de um campeonato de futebol, um grande show de rock, a véspera de Ano Novo, a realização dos Jogos Olímpicos ou os jogos de uma Copa do Mundo. Mesmo se forem conhecidos o público esperado, o horário e a data do evento, é comum perceber o congestionamento dos recursos das operadoras de telefonia celular durante eventos dessa proporção, decorrente de mudanças coletivas na carga de trabalho da rede (BAGROW; WANG; BARABáSI, 2011).

Como é de amplo conhecimento, o Brasil sediou a Copa do Mundo de Futebol em 2014 e receberá os Jogos Olímpicos em 2016. Segundo dados da Polícia Federal, aproxi-

madamente 700 mil turistas, de todos os lugares do mundo, entraram no país durante o mês de junho de 2014 e outros milhares concentrarão a visita ao nosso pais no período de realização das Olimpíadas de 2016. Para atender as necessidades dessas pessoas, interessadas em um evento de larga escala, investimentos em setores como mobilidade e infraestrutura são necessários.

Pesquisas que utilizam dados gerados pelos usuários do serviço de telefonia móvel, mesmo que ainda escassas no Brasil, podem auxiliar a identificar locais e regiões onde a infraestrutura atual não suporta a demanda de serviços. O estudo desses dados possibilita também a análise de como será o acesso aos locais do evento, como os serviços de telefonia serão utilizados, para qual local que os participantes do evento irão após seu término, qual o horário que os participantes começam a ir para o local do evento, entre outras informações. Uma terceira possibilidade da análise desses dados permite compreender os padrões de comportamento da mobilidade humana, identificados a partir do histórico de uso de serviços de telefonia móvel. Compreender a mobilidade pode aumentar a precisão da caracterização dos usuários, visando apoiar o desenvolvimento de melhores estratégias de gestão de recursos e serviços.

Entender qual o horário os participantes de um evento de larga escala começam a ir para o local de sua realização e para qual local irão após seu término, envolve diretamente uma análise da mobilidade dos mesmos. O estudo de Song et al. (2010b) mostra que a mobilidade humana pode ser prevista, com razoável precisão, a partir da análise das rotinas diárias das pessoas. Contudo, compreender os padrões de movimento das pessoas, utilizando um conjunto de dados gerados pelo uso de telefones móveis, durante uma curta janela de tempo é uma tarefa muito complexa, pois, a análise depente da realização de pelo menos uma chamada durante essa janela de tempo.

As informações mantidas nos registros gerados pela utilização dos serviços de telefonia móvel como estações bases utilizadas, duração da chamada e horário da chamada, por exemplo, podem fomentar a criação de modelos científicos que auxiliem o estudo da mobilidade e consequentemente o planejamento eficiente da rede, obviamente mantendo o anonimato do usuário. Os modelos criados podem, consequentemente, se transformar em ferramentas que visem agilizar a execução dessas análises. Estudos que desenvolvam ferramentas como essas ainda são raros, se mostrando assim uma oportunidade para a área de Ciência da Computação.

Este trabalho analisa características da rede de telefonia móvel e a mobilidade humana durante alguns tipos de eventos de larga escala, especificamente partidas de futebol, Réveillon e grandes shows. Estas análises utilizam de bases de dados referentes a chamadas de voz de uma grande operadora brasileira. As informações presentes nessas bases de dados são totalmente anônimas, garantido assim a privacidade dos usuários.

Para o desenvolvimento destas análises foi desenvolvida uma metodologia específica para grandes eventos que também será apresentada neste trabalho.

O trabalho de caracterização utiliza uma metodologia de análise aplicada a bases de dados contendo as chamadas feitas durante jogos do Campeonato Brasileiro, Campeonato Mineiro, Copa do Brasil, Amistosos da Seleção Brasileira, Copa das Confederações e Copa do Mundo. Além dos jogos de futebol também foram analisadas as comemorações do Ano Novo em Belo Horizonte, Rio de Janeiro, São Paulo, Recife e Salvador e shows em Belo Horizonte e Rio de Janeiro. Foi analisada também uma base de dados contendo informações sobre o Grande Prêmio de Interlagos, totalizando assim 29 bases de dados analisadas.

A caracterização da carga de trabalho é analisada em termos da quantidade, tempos entre início e término de chamadas sucessivas, e duração das chamadas feitas pelos usuários que estiveram presentes na região do evento. Este trabalho mostra a evolução de uma metodologia de caracterização e apresenta uma metodologia para análise da mobilidade humana em grandes eventos. Os resultados da aplicação das metodologias são apresentados inicialmente com uma análise exclusiva de cada tipo de evento, progredindo para uma análise comparativa da região onde o mesmo é realizado e sendo finalizados com o estudo da mobilidade dos usuários.

Além do trabalho de caracterização também foi desenvolvida uma ferramenta, denominada de *PrediTraf*. Esta ferramenta realiza a análise e previsão da dinâmica da carga de trabalho e também da mobilidade dos usuários. Todas as análises realizadas pelo *PrediTraf* utilizam de eventos anteriores visando garantir uma maior precisão nas previsões de eventos futuros.

O restante deste trabalho está organizado como segue. O Capítulo 2 discute trabalhos relacionados. As bases de dados utilizadas e a metodologia de análise adotada são apresentadas nos Capítulos 3 e 4, respectivamente. Os principais resultados das análises realizada são discutidos no Capítulo 5. A ferramenta *PrediTraf* é apresentada no Capítulo 6. Conclusão e trabalhos futuros são apresentados no Capítulo 7.

2 REVISÃO BIBLIOGRÁFICA

2.1 Caracterização de Carga de Trabalho

Com um crescimento anual de aproximadamente 40% no número de usuários dos serviços de telefonia móvel em todo o mundo, segundo Ericsson (2014), se torna cada vez mais necessário para as operadoras compreenderem o comportamento dos seus usuários para atender a demanda da rede e oferecer serviços de qualidade. Para compreender a carga de trabalho gerada em uma rede móvel, pela realização de um grande evento, é importante realizar uma comparação do uso da rede em dias comuns e dias com eventos (PULSELLI et al., 2008). Essa comparação permite criar modelos matemáticos que ajudem a compreender a utilização dos recursos em dias com e sem eventos (SONG et al., 2010b). Percebe-se que o tipo de carga de trabalho varia de acordo com a natureza do evento (show, acidente, jogos de futebol, etc.) possuindo assim comportamento específico de acordo com esta natureza (BAGROW; WANG; BARABáSI, 2011).

No trabalho de Pulselli et al. (2008) é realizada uma análise durante dezesseis dias da cidade de Milão, afim de verificar o uso da rede relacionando-o com a ocupação física do local. Baseando-se no estudo de Pulselli et al. (2008), em Calabrese et al. (2010) é realizada uma análise de diversos eventos, de natureza e público diferentes. A partir deste trabalho também é possível afirmar que eventos do mesmo tipo atraem participantes de regiões específicas da cidade. Conhecendo o comportamento de públicos distintos e suas origens dentro da cidade, se pode desenvolver mecanismos de melhoria da mobilidade urbana, afim de evitar congestionamentos das principais vias de acesso ao evento (CALABRESE et al., 2010).

Um problema comum na análise de grandes eventos está na detecção de áreas onde ocorre uma aglomeração de usuários. Identificar uma região geográfica específica e durante um período de tempo, como em eventos de larga escala, ainda é um desafio (MORLOT; ELAYOUBI; BACCELLI, 2010). As operadoras de telefonia celular normalmente encontram dificuldades para planejar a infraestrutura necessária para a utilização de serviços, em situações onde o uso da rede gera cargas fora do padrão (BAGROW; WANG; BARABáSI, 2011).

Para compreender e modelar o padrão de carga gerado por uma rede de telefonia móvel, Filiposka e Mishkovski (2013) desenvolveram um gerador de tráfego móvel capaz de produzir uma carga sintética de voz, SMS e tráfego de dados em uma rede, baseado em uma carga real. Uma característica importante ao se desenvolver este simulador de

redes móveis, está no fato de que o ambiente de simulação deve, tanto quanto possível, refletir um cenário real. Essa simulação deve ser independente da infraestrutura utilizada e tecnologia empregada na rede existente. Após combinarem dados reais, Filiposka e Mishkovski (2013) conseguiram obter uma carga sintética de trafego que representa, de forma satisfatória, o comportamento original analisado. Os autores verificaram porém que ainda é difícil conseguir sintetizar tais dados que se assemelhem ao tráfego real.

Trabalhos como o de Filiposka e Mishkovski (2013) motivaram o desenvolvimento da ferramenta *PrediTraf* . A possibilidade de desenvolver uma aplicação que implementa a metodologia apresentada neste trabalho se mostrou viável e necessária. Analisar novos eventos baseando-se em eventos anteriores, além de utilizar dos dados gerados pela execução da metodologia de caracterização é uma oportunidade de validação da metodologia e desenvolvimento de uma ferramenta inovadora.

2.2 Mobilidade Humana

Além de caracterizar a carga de trabalho de eventos de natureza distintas, este trabalho também analisa a mobilidade dos usuários presentes na região do evento. A mobilidade humana pode ser descrita como uma rede de interações entre comunidades, onde as conexões e a intensidade representam respectivamente o fluxo de pessoas e as iterações entre os membros (BALCAN et al., 2009). A mobilidade humana é dividida em três dimensões: espacial (e.g. distância percorrida por usuário), temporal (e.g. frequência de visitas à um local) e social (e.g. contato entre usuários) como mostrado em Asgari, Gauthier e Becker (2013) e (KARAMSHUK et al., 2011).

Para compreender a mobilidade humana, vários estudos têm analisado diversas fontes de informação tentando compreender os padrões de locomoção e localização do ser humano durante suas atividades diárias. Esses podem utilizar de diversas fontes de informação como a circulação de notas de banco mostrada em Brockmann, Hufnagel e Geisel (2006). Jiang e Jia (2011) apresentou o uso de US Flights, Lane et al. (2010) utilizou de dados de GPS e os registros de chamadas (CDRs) foi o objeto de estudo em (FICEK; POP; KENCL, 2013).

Como dito na Seção 2.1, um problema comum na análise de eventos de larga escala está na detecção de áreas onde ocorre uma aglomeração de usuários. Para resolver esses problemas a modelagem dos padrões de localização humana vem sendo analisada por vários estudos e propondo novas direções. Esse grande número de estudos se deve ao aumento da disponibilidade de grandes fontes de dados da atividade humana. Trabalhos como Hasan et al. (2013) apresentam uma ideia de seleção preferencial aos locais de interesse (trabalho, escola, casa) e pontos turísticos da cidade. Essa ideia permite modelar a distribuição espacial dos movimentos do indivíduo, dentro de um centro urbano.

Para isso, os autores utilizam de dados obtidos a partir de cartões de passagem de metrô em grandes cidades, além de uma seleção de locais não-residenciais e não-comerciais. O resultado deste trabalho mostrou que é possível observar que os locais visitados com maior frequência podem ser modelados com probabilidades fixas de visitas. Desta maneira comprova-se que os usuários tendem a ir aos locais considerados mais interessantes, consequentemente com maior probabilidade, da cidade quando escolhem seus destinos.

O principal dado utilizado para estudos da mobilidade dos usuários de telefonia móvel são seus registros de chamadas, os CDR (Call Data Record). Uma vez que telefones móveis são transportados durante a rotina diária, esta característica facilita capturar as trajetórias de cada indivíduo (GONZALEZ; HIDALGO; BARABASI, 2008). Esses registros são mantidos pelas operadoras de serviço de telefonia móvel, podendo ser analisados para compreender e prever os padrões de movimentação humana (de onde as pessoas saem, e para onde vão) durante o dia. Estudos mostram que utilizando desses dados é possível encontrar, com até 93% de precisão, a localização do usuário durante sua rotina diária (SONG et al., 2010b).

A utilização de CDRs para trabalhos de pesquisa e análise do comportamento humano requerem uma abordagem holística e cuidadosa sobre privacidade dos dados dos usuários, devido à gama de aplicações que esses possibilitam (LAURILA et al., 2012). Um exemplo dessas aplicações são os estudos dos registros de chamadas realizados pelos usuários de telefonia móvel não apenas quanto à mobilidade humana mas para também extrair informações relevantes acerca da estrutura das cidades (LOUAIL et al., 2014). A partir da análise de CDRs, é possível inferir a dinâmica de utilização dos locais de uma cidade (área industrial, residencial, comercial, etc.), a fim de compreender a heterogeneidade da utilização dos espaços feitos pela população de forma temporal (TOOLE et al., 2012; YUAN; ZHENG; XIE, 2012).

Em Cho, Myers e Leskovec (2011) inferiu-se que padrões de mobilidade estão fortemente relacionados às relações sociais dos indivíduos. A influência da amizade na mobilidade de um usuário pode ser duas vezes mais forte que a influência da mobilidade sobre a criação de novas amizades. Em Isaacman et al. (2011), os autores demostram que os locais onde os usuários de telefonia móvel visitam, mesmo que realizem poucas chamadas, podem ser encontrados com cerca de 88% de exatidão num limite de 3 milhas. A utilização dos aparelhos móveis pode possibilitar a determinação da amizade entre os usuários e consequentemente auxiliar o estudo de sua mobilidade.

Estudo de Hohwald, Frías-Martínez e Oliver (2010) confirmou que os padrões de comunicação entre os usuários dependem da hora e do dia da semana. Notou-se que os usuários de telefonia móvel são mais susceptíveis a se comunicar com pessoas de seu círculo social. Em Hidalgo e Rodriguez-Sickert (2008) é apresentado que as chamadas

de celular podem servir como identificador de laços sociais. Os autores concluíram que amigos tendem a realizar chamadas entre si frequentemente. Dessa maneira, os dados obtidos a partir de ligações telefônicas possuem o potencial de prover informações sobre a dinâmica relacional de comunidades e sobre as características de utilização da rede de telefonia celular pelos usuários (EAGLE; PENTLAND; LAZER, 2009).

Uma forma de compreender o comportamento humano é através da análise de seus padrões de movimentação. Em Gonzalez, Hidalgo e Barabasi (2008), os autores utilizam de modelos matemáticos para identificar padrões de movimentação diários dos usuários da rede móvel, analisando as rotas utilizadas e principais locais de aglomeração. As rotinas diárias de movimentação dos usuários são de grande importância para compreender alterações na rede (WANG et al., 2009). Esses estudos podem possibilitar o relacionamento dos padrões de aglomeração à geometria do ambiente de determinados eventos (BATTY; DESYLLAS; DUXBURY, 2003).

Em Song et al. (2010a) os autores modelam o padrão de mobilidade do ser humano através de trajetórias individuais coletadas. Esse trabalho permitiu demonstrar a partir de uma série de medidas, que as trajetórias humanas seguem leis de escala altamente reprodutíveis. Em Halepovic e Williamson (2005) por sua vez, os autores utilizam o tráfego de dados em uma rede e a concentração de utilização da rede ao longo do tempo, para desenvolverem um modelo de mobilidade. O modelo desenvolvido permite os autores simularem as propriedades de mobilidade dos usuários na rede celular. O trabalho de Gonzalez (2013) compara dados de mobilidade entre duas cidades diferentes, e revelam que os padrões de mobilidade humana não se diferem muito entre si. Esse estudo mostra ainda que os dados podem ser usados para modelar e simular a atividade urbana, permitindo assim auxílio a tomada de decisões nas cidades.

A predição da mobilidade humana pode ajudar a melhorar serviços destinados à sociedade. O trabaho de Wang et al. (2010) utiliza registros de chamadas de telefonia móvel e propõem um método para inferir o meio de transporte utilizado pelos usuários da rede. A compreensão de tal padrão de mobilidade é de grande importância para o melhor gerenciamento e planejamento dos sistemas de transporte em grandes cidades. Em Frias-Martinez, Soguero e Frias-Martinez (2012) a partir dos dados de utilização de telefonia celular, os autores conseguem determinar padrões urbanos de mobilidade e a utilização de espaço. Os estudos apresentaram um alto grau de similaridade aos dados obtidos a partir de pesquisas e questionários realizados por Institutos Nacionais de Pesquisas, como o *United States Census Bareau*.

Em Jiang et al. (2013) os autores identificam as paradas (i.e. "*stays*", chamadas realizadas enquanto o usuário está realizando alguma atividade sem se movimentar), e as caminhadas (i.e. "*pass-by*", chamadas realizadas enquanto os usuários se movimentam).

Nesse trabalho os autores analisam a mobilidade humana e derivam características universais que revelam a previsibilidade da mobilidade. O estudo mostra que esta mobilidade segue padrões de movimentações. Ainda neste estudo, apresentam uma classe de algoritmos para extrair áreas onde há uma grande concentração de permanência e caminhadas de cada indivíduo, além de locais de estadia em potencial. Os algoritmos descritos no trabalho apresentaram um grau de acerto próximo de 70%.

O estudo da mobilidade humana apresenta diversas maneiras de serem realizados e possibilitam a descoberta de conhecimentos de diversas áreas. Entretanto, o trabalho aqui desenvolvido tem como objetivo a análise da mobilidade humana específica em grandes eventos assim como Calabrese et al. (2010) apresentou. Neste trabalho o autor mostra que existe uma forte correlação entre pessoas que vivem próximas aos eventos de larga escala serem preferencialmente atraídos por eles. Os autores concluem também que eventos da mesma natureza apresentam uma distribuição similar da origem dos participantes. Como consequência, poderiam predizer de onde as pessoas tendem a vir de eventos futuros, facilitando o entendimento da movimentação das pessoas dentro deste espaço. Desta maneira o estudo dos CDRs pode caracterizar de forma precisa vários aspectos da mobilidade humana (ISAACMAN et al., 2012; YAN CHEN ZHAO, 2014).

2.3 Uso de Dados de Telefonia para Construção de Cidades Inteligentes

Os resultados da caracterização da carga de trabalho da rede de telefonia móvel concomitante com o estudo da mobilidade humana podem auxiliar a criação de cidades inteligentes. As cidades estão se tornando inteligentes não apenas nos aspectos, de como podemos automatizar rotinas para utilização das pessoas, edificações e sistemas de tráfego, mas, também de maneira que nos permitam monitorar, compreender e analisar o comportamento do ser humano afim de melhorar a eficiência e qualidade de vida dos usuários nos grandes centros urbanos (BATTY et al., 2012).

Uma das linhas de pesquisa consiste em analisar dados acerca do comportamento humano. Os potenciais benefícios de analisar tais dados são vários. Esses benefícios vão desde a diminuição de tráfego até a uma melhor utilização de recursos públicos como parques, transporte coletivo e serviço de ambulâncias. Tais estudos podem providenciar também informações capazes de auxiliar uma melhoria dos serviços prestados a sociedade, como os serviços de telecom (MITCHELL, 2009). O controle de doenças também pode ser um benéficio da análise dos dados do comportamento humano. A disseminação de doenças em uma cidade ocorre muitas vezes através das iterações entre as pessoas. As ligações telefônicas podem ser utilzadas para compreender padrões de movimento e consequentemente a forma de contágio de diversas doenças (BALCAN et al., 2009) auxiliando assim o planejamento de cidades inteligentes.

Uma alternativa à instalação de sensores específicos para monitorar o fluxo de veículos em uma cidade consiste na utilização da rede de telefonia celular para auxiliar tal monitoramento. Estudos da rede de telefonia celular correlacionada ao fluxo de tráfego de veículos nas vias podem determinar além da quantidade de automóveis que transitam na via também a relação origem-destino dos usuários (CACERES; WIDEBERG; BENITEZ, 2007). Os resultados do trabalho aqui apresentado auxiliam a compreensão da mobilidade em situações fora da rotina diária, gerando assim subsídios que permitam uma tomada de decisão automatizada em uma cidade.

Padrões anômalos, como cargas fora do padrão podem indicar alguma crise (e.g. catástrofe natural, ataque terrorista, etc.) ou algum evento fora de rotina como concertos musicais, jogos de futebol, entre outros (CANDIA et al., 2008). O padrão de movimentação e utilização da rede de telefonia móvel durante uma emergência é alterado consideravelmente. Estudos mostram que esses eventos causam um aumento instantâneo no número de chamadas logo após o ocorrido e com o decorrer do tempo uma diminuição gradativa. Este comportamento indica a necessidade dos usuários de informar imediatamente as suas interações (GAO et al., 2014). A metodologia aqui apresentada permite uma análise de períodos próximos ao horário do evento. Caso realizada em tempo real pode-se encontrar pontos onde estejam acontecendo situações críticas, como um confronto entre torcidas rivais. Além disso, ferramentas como a desenvolvida neste trabalho podem auxiliar governos a promover uma análise da situação de suas cidades e, consequentemente, auxiliar tomadas de decisões que as tornem cidades cada vez mais inteligentes.

3 BASES DE DADOS

3.1 Estrutura da Base de Dados

A metodologia que será apresentada no Capítulo 4 foi aplicada a diversas bases de dados disponibilizadas por uma grande operadora de telefonia móvel brasileira. Essas bases foram adquiridas a partir da parceria para desenvolvimento tecnológico firmada em 2012 entre a PUC MINAS e a Oi S/A. Todos os dados aqui analisados foram coletados a partir da ferramenta Traffica*. Cada base de dados contém informações sobre as chamadas realizadas a partir de telefones celulares em uma determinada cidade e em um dado dia, note que as bases de dados coletadas representam apenas uma amostra do uso da rede de telefonia móvel. Para cada chamada, os registros disponilizados contêm os seguintes campos:

- Identificador único associado à chamada;
- Identificador único associado ao usuário, gerado pela operadora e completamente anonimizado; Este campo não leva à identificação do usuário mas, permite identificar todas as chamadas feitas pelo mesmo usuário;
- Horários de início e fim de cada chamada;
- Identificadores das Estações Rádio Bases (ERBs) onde a chamada foi iniciada e finalizada.

A operadora de telefonia disponibilizou também a base de dados de suas ERBs referentes as cidades analisadas. A base de dados disponibilizada contém a latitude e longitude de cada ERB possibilitando, assim, o georreferenciamento dos usuários utilizando. Note que a cobertura de cada ERB é dividida em regiões denominadas de setores, que possuem os seguintes campos:

- Identificador único da ERB;
- Identificador único do setor, que permite a identificação do local onde a chamada foi iniciada ou finalizada;

*http://networks.nokia.com/portfolio/products/customer-experience-management/serve-atonce-traffica

- Localizações geográficas (latitude e longitude) das ERBs;
- Modelo da ERB;
- Número de rádios transmissores e receptores (TRXs);

3.2 Eventos Coletados

Durante o desenvolvimento da pesquisa foram coletados dados da rede de telefonia móvel de seis cidades diferentes: Rio de Janeiro, São Paulo, Belo Horizonte, Sete Lagoas, Salvador e Recife. Destas coletas foram selecionadas para este trabalho as que melhor retratavam o comportamento dos usuários, além de priorizar aquelas que ainda não haviam sido apresentadas em trabalhos anteriormente publicados pelo autor. Considerou-se bases que retratam o comportamento dos usuários aquelas que apresentaram um número de chamadas compatível com o público presente do evento de larga escala, caso este número não tenha sido atingido a base de dados foi descartada.

Essa seleção das bases de dados foi realizada para uniformização das análises e conslusões acerca do comportamento dos usuários. Como as chamadas são o principal objeto de estudo, bases de dados que apresentaram problemas devido a diversos fatores, entre eles uma falha de coleta, não foram utilizadas. Por exemplo, o show realizado pela cantora Madonna no Morumbi em São Paulo é uma dessas bases que não foram utilizadas para a caracterização dos eventos de larga escala. O público deste evento realizado em 02 de dezembro de 2012 foi de 58 mil pessoas, entretanto apenas 89 chamadas foram coletadas. Este pequeno valor se mostrou insuficiente para o trabalho pois nenhuma consideração poderia ser feita a partir deste conjunto reduzido de chamadas. Contudo, note que todas as coletas realizadas estarão listadas em anexo ao trabalho. Os eventos analisados foram divididos em 4 grupos, (I) Partidas de Futebol, (II) Shows, (III) Réveillon e (IV) Automobilismo. Os eventos de larga escala coletados entre os anos de 2011 e 2014 utilizados neste trabalho serão descritos a seguir.

3.2.1 Partidas de Futebol

As informações da rede de telefonia móvel foram coletadas durante as partidas de futebol nas cidades de Belo Horizonte, Fortaleza, Recife e Rio de Janeiro, em datas específicas durante os anos de 2011 a 2014. Foram coletados dados da rede de telefonia móvel em jogos nos estádios Mineirão e Independência, em Belo Horizonte, Arena Castelão em Fortaleza, Arena Pernambuco em Recife, Maracanã e Engenhão no Rio de Janeiro. A Tabela 1 mostra as informações sobre os eventos analisados.

Tabela 1 – Informações das partidas de futebol analisadas

Cidade	Evento	Data	Horário	Público Presente	Local
Belo Horizonte	Atlético x Fluminense	21/10/2012	16:00	20.096	Independência
	Cruzeiro x Atlético	03/02/2013	17:00	52.989	Mineirão
	Brasil x Chile	24/04/2013	22:00	53.331	Mineirão
Rio de Janeiro	Flamengo x Vasco	28/08/2011	16:00	33.206	Engenhão
	Flamengo x Vasco	04/12/2011	17:00	40.004	Engenhão
	Flamengo x Fluminense	08/07/2012	16:00	32.591	Engenhão
	Brasil x Espanha	30/06/2013	19:00	73.531	Maracanã
	Flamengo x Atlético PR	27/11/2013	22:00	68.857	Maracanã
	Colômbia x Uruguai	28/06/2014	17:00	73.804	Maracanã
	Alemanha x Argentina	13/07/2014	16:00	74.738	Maracanã
Pernambuco	Grécia x Costa Rica	29/06/2014	17:00	41.242	Arena PE
Fortaleza	Holanda x México	29/06/2014	13:00	58.817	Castelão

Em Belo Horizonte, foram coletadas informações da rede de telefonia móvel durante três jogos, o primeiro jogo ocorreu em 21 de outubro de 2012, partida de futebol válida pelo Campeonato Brasileiro entre o Clube Atlético Mineiro e Fluminense Football Club, no estádio Independência. Essa partida de futebol iniciou às 16 horas e teve um público de 20.096 espectadores. Uma característica importante desse jogo é que o mesmo era tratado como uma final, pois ambos disputavam o título do campeonato em 2012. A coleta das chamadas dessa partida de futebol iniciou-se às 12 horas e terminou às 21 horas.

Os outros dois jogos de Belo Horizonte ocorreram no estádio Mineirão. O primeiro jogo ocorreu em 03 de fevereiro de 2013 entre o Cruzeiro Esporte Clube e o Clube Atlético Mineiro, as duas principais equipes da cidade. Partida de futebol válida pelo Campeonato Mineiro de 2013, marcou também a reinauguração do estádio que ficou fechado durante três anos para reforma. Essa partida de futebol contou com um público de 52.989 espectadores e iniciou-se às 17 horas. A coleta dos dados da rede de telefonia móvel iniciou-se às 12 horas e foi finalizada às 20 horas.

O segundo jogo coletado no estádio Mineirão, na cidade de Belo Horizonte, foi um amistoso internacional entre as seleções do Brasil e Chile no dia 24 de abril de 2013. O número de espectadores da partida de futebol foi de 53.331 e a mesma iniciou às 22 horas. Para esta partida de futebol a coleta dos dados da rede de telefonia móvel iniciou-se às 18 horas do dia 24 de abril e terminou às 02 horas do dia 25 de abril.

A cidade de Fortaleza recebeu a partida de futebol entre Holanda e México válido pelas oitavas de final da Copa do Mundo de 2014. Esta partida de futebol ocorreu no dia 29 de junho de 2014 sendo iniciada às 13 horas. Estiveram presentes 58.817 espectadores. A coleta dos dados da rede de telefonia móvel desse evento foi iniciada às 12 horas e finalizada às 20 horas.

A partida de futebol entre Grécia e Costa Rica, também válida pelas oitavas de final da Copa do Mundo de 2014, foi disputada na Arena Pernambuco em Recife. Iniciada

às 17 horas, a partida de futebol contou com a presença de 41.242 espectadores. Os dados da rede de telefonia móvel foram coletados a partir das 13 horas e 30 minutos sendo a coleta finalizada às 21 horas e 30 minutos.

Na cidade do Rio de Janeiro foram coletados dados da rede de telefonia móvel durante a realização de sete partidas de futebol, sendo quatro no estádio Maracanã, em datas específicas nos anos de 2013 e 2014, e as demais no estádio Engenhão em datas específicas nos anos de 2011 e 2012. Os dados da rede de telefonia móvel das duas primeiras partidas de futebol, na cidade do Rio de Janeiro, foram coletados em 2011 no estádio Engenhão. As partidas de futebol foram disputadas entre o Clube de Regatas do Flamengo e Club de Regatas Vasco da Gama. Ambas as partidas de futebol foram válidas pelo Campeonato Brasileiro de 2011. Em 28 de agosto de 2011 ocorreu o primeiro jogo entre as duas equipes, esta partida de futebol teve um público de 33.206 espectadores iniciando às 16 horas. A coleta dos dados da rede de telefonia móvel durante esta partida de futebol iniciou-se às 13 horas e foi finalizada às 20 horas.

A segunda partida de futebol, na cidade do Rio de Janeiro, ocorreu em 04 de dezembro de 2011. Em caso de vitória da equipe do Flamengo resultaria na conquista do título de campeão brasileiro de 2011. Em virtude do horário de verão brasileiro esta partida de futebol ocorreu às 17 horas e contou com um público de 40.004 espectadores. A coleta dos dados da rede de telefonia móvel foi iniciada às 09 horas e 30 minutos e finalizada às 21 horas.

Em 08 de julho de 2012 foram coletados dados da rede de telefonia móvel referente à partida de futebol entre o Clube de Regatas do Flamengo e Fluminense Football Club, um dos principais clássicos da cidade do Rio de Janeiro. Esta partida de futebol ocorreu no estádio Engenhão iniciando às 16 horas e foi acompanhada por 32.591 espectadores. A coleta dos dados da rede de telefonia móvel foi realizada a partir das 13 horas e 30 minutos e finalizada às 20 horas.

A coleta dos dados da rede de telefonia móvel do primeiro jogo no estádio Maracanã, na cidade do Rio de Janeiro, aconteceu em 30 de junho de 2013 durante a partida de futebol entre as seleções de Brasil e Espanha. Este jogo marcou a final das Copa das Confederações e contou com a presença de 73.531 espectadores. Os dados da rede de telefonia móvel referente a essa partida de futebol foram coletados a partir 14 horas e 30 minutos até às 23 horas.

A coleta dos dados da rede de telefonia móvel do segundo jogo no estádio Maracanã aconteceu na data de 27 de novembro de 2013 às 22 horas entre o Clube de Regatas do Flamengo e o Clube Atlético Paranaense. Esta partida de futebol era válida como o segundo jogo da final da Copa do Brasil daquele ano. O público presente foi de 68.857 espectadores. O resultado desta partida de futebol levou a equipe do Flamengo ao título

de campeão da Copa do Brasil de 2013. Para essa partida de futebol a coleta dos dados da rede de telefonia móvel foi iniciada às 17 horas e 30 minutos e finalizada à zero hora do dia 28 de novembro.

As duas coletas dos dados da rede de telefonia móvel mais recentes na cidade do Rio de Janeiro foram realizadas no estádio Maracanã durante a Copa do Mundo de 2014. Foram coletados dados referentes as partidas de futebol da oitava de final e da final em 28 de junho às 17 horas e 13 de julho às 16 horas respectivamente.

A partida referente às oitavas de final ocorreu entre as seleções da Colômbia e Uruguai. Neste evento a coleta dos dados da rede de telefonia móvel iniciou-se às 13 horas e 30 minutos sendo finalizada às 21 horas e 30 minutos. Como é de amplo conhecimento, a final da copa do mundo ocorreu entre as seleções da Argentina e Alemanha. A coleta dos dados da rede de telefonia móvel desta partida de futebol iniciou-se às 12 horas e 50 minutos e foi finalizada às 20 horas e 50 minutos.

Afim de realizar uma análise sobre o impacto que um evento de larga escala causa na região onde é realizado, também foram coletados dados da rede de telefonia móvel na região do estádio Engenhão quando o mesmo não recebia nenhuma partida de futebol ou show. A data de 11 de dezembro de 2011 foi escolhida para a coleta dos dados da rede de telefonia móvel na região. A coleta desses dados iniciou-se ao meio dia sendo finalizada às 20 horas e 30 minutos.

3.2.2 Shows

As coletas dos dados de telefonia móvel durante shows foram realizadas apenas nas cidades de Belo Horizonte e Rio de Janeiro. Foram realizadas apenas três coletas desses dados sendo duas em Belo Horizonte no ano de 2013 e uma no Rio de Janeiro no ano de 2012. A Tabela 2 mostra as informações sobre os eventos analisados.

Tabela 2 – Informações sobre os shows analisados

Cidade	Evento	Data	Horário	Público	Local
Belo Horizonte	Show Paul McCartney	04/05/2013	21:30	53.000	Mineirão
	Show Beyoncé	11/09/2013	21:30	30.000	Mineirão
Rio de Janeiro	Show Roger Waters	29/03/2012	21:30	50.000	Engenhão

Em Belo Horizonte as duas coletas dos dados da rede de telefonia móvel foram realizadas no estádio Mineirão e os shows iniciaram às 21 horas e 30 minutos. O primeiro evento coletado aconteceu em 04 de maio de 2013. Nesta data o músico Sir James Paul McCartney realizou sua apresentação durante 2 horas. Sendo um dos criadores de umas das principais bandas de rock da história, The Beatles, gerou-se uma grande expectativa

para o show que atraiu 53.000 espectadores ao estádio. A coleta dos dados da rede de telefonia móvel durante esse evento iniciou-se às 16 horas do dia 04 de maio e foi finalizada às 2 horas do dia 05 de maio.

O segundo evento em Belo Horizonte ocorreu em 11 de setembro de 2013. Nesta data 30.000 espectadores estiveram presentes ao estádio Mineirão para assistir à apresentação da cantora Beyoncé Knowles. Assim como o show de Paul McCartney a duração deste evento foi de 2 horas e iniciou às 21:30. Neste evento a coleta dos dados da rede de telefonia móvel foi iniciada às 17 horas do dia 11 de setembro e finalizada à meia noite do dia 12 de setembro.

O evento coletado na cidade do Rio de Janeiro ocorreu no estádio Engenhão na data de 29 de março de 2012 às 21 horas e 30 minutos. Nesta data ocorreu o show do músico George Roger Waters, um dos fundadores da famosa banda Pink Floyd. Este evento atraiu 50.000 espectadores e também teve duração de 2 horas. Os dados da rede de telefonia móvel durante este evento foram coletados a partir das 17 horas do dia 29 de março tendo a coleta finalizada à 1 hora da manhã do dia 30 de março.

3.2.3 Réveillon

As coletas dos dados da rede de telefonia móvel durante as festas de Réveillon ocorreram durante a virada do ano de 2011 para 2012 nas cidades de Belo Horizonte, Recife, Rio de Janeiro, São Paulo e Salvador. Os eventos coletados permitem analisar festas de duas características diferentes, sendo elas públicas e privadas. O clube Cabanga, localizado na cidade de Recife, recebeu 10.000 espectadores para a festa de Réveillon, ao contrário dos demais Réveillons analisados a entrada para esse evento não foi pública. Nas demais cidades as comemorações do Réveillon foram abertas ao público. A Tabela 3 mostra as informações sobre os Réveillons analisados.

Tabela 3 – Informações sobre os Réveillons analisados

Cidade	Data	Público	Local
Belo Horizonte	31/12/2011	100.000	Praça da Estação
Rio de Janeiro	31/12/2011	2,5 Milhões	Praia de Copacabana
Recife	31/12/2011	10.000	Clube Cabanga
Salvador	31/12/2011	1 Milhão	Praia da Barra
São Paulo	31/12/2011	2 Milhões	Avenida Paulista

Em Belo Horizonte a comemoração da virada do ano aconteceu na Praça Rui Barbosa, também conhecida como Praça da Estação. Neste ano a festa foi patrocinada por uma emissora de televisão e contou com a presença de 150.000 espectadores, segundo a Polícia Militar. Na cidade de Salvador a festa de Réveillon analisada ocorreu na Praia

da Barra, considerado uns dos mais famosos Réveillons do nordeste brasileiro. Estima-se que 1 milhão de espectadores estavam presentes nesta festa no dia 31 de dezembro.

A maior comemoração do Réveillon analisada ocorreu na cidade do Rio de Janeiro, mais especificamente na Praia de Copacabana. Sendo a principal festa de Réveillon do Brasil, a Praia de Copacabana recebe turistas do mundo inteiro. Estima-se que cerca de 2,3 milhões de espectadores estiveram presentes nessa data e local.

Em São Paulo, a festa de Réveillon aconteceu na Avenida Paulista. A Polícia Militar estima que no ano de 2011 aproximadamente 2 milhões de espectadores estiveram presentes na festa. A coleta dos dados da rede de telefonia móvel durante o Réveillon de São Paulo iniciou-se as 19 horas do dia 31 de dezembro de 2011 e foi finalizada as 04 horas da manhã do dia 1 de Janeiro de 2012. Ressalta-se que este período de coleta também foi adotado para as demais festas de Réveillon.

Assim como no estádio Engenhão, para as festas de Réveillon foram coletados dados da rede de telefonia móvel na região quando a mesma não recebia nenhum evento. Para todas as regiões a data escolhida foi entre os dias 03 de Janeiro de 2012 e 04 de Janeiro de 2012. A coleta dos dados da rede de telefonia móvel na cidade do Rio de Janeiro iniciou-se as 21 horas do dia 03 e foi finalizada as 03 horas e 30 minutos do dia 04. Para as demais cidades as coletas dos dados da rede de telefonia móvel foram iniciadas as 19 horas do dia 03 e finalizadas as 04 horas do dia 04 de janeiro.

3.2.4 Automobilismo

O único evento de automobilismo analisado aconteceu em 25 de novembro de 2012 às 14 horas no autódromo de Interlagos em São Paulo. Esse evento marcava o final da temporada de Fórmula 1 do corrente ano e o público presente foi de 69.984 espectadores. Nessa data ainda estava aberta a disputa do título entre Fernando Alonso e Sebastian Vettel, sendo confirmado para o segundo. A corrida foi vencida pelo piloto Jenson Button e o piloto brasileiro Felipe Massa chegou em 3º lugar. A coleta dos dados da rede de telefonia móvel durante esse evento foi iniciada às 11 horas e finalizada às 18 horas do dia 25 de novembro. Assim como nos demais eventos foi realizada uma coleta da rede de telefonia móvel da região de Interlagos quando não havia nenhum evento sendo realizado. Essa coleta aconteceu no dia 24 de março de 2013 sendo iniciada às 12 horas e finalizada as 20 horas. A Tabela 4 apresenta as informações do evento.

Tabela 4 – Corridas de Automóveis coletadas

Cidade	Evento	Data	Horário	Público	Local
São Paulo	Grande Prêmio do Brasil de F1	25/11/2012	14:00	69.984	Autódromo de Interlagos

4 METODOLOGIA

A metodologia proposta neste trabalho pode ser dividida em duas partes, sendo a primeira responsável pela caracterização do usuário (metodologia de caracterização) e a segunda pela análise detalhada da mobilidade humana (metodologia de mobilidade). A Figura 1 apresenta os principais objetivos de cada parte da metodologia proposta neste trabalho.

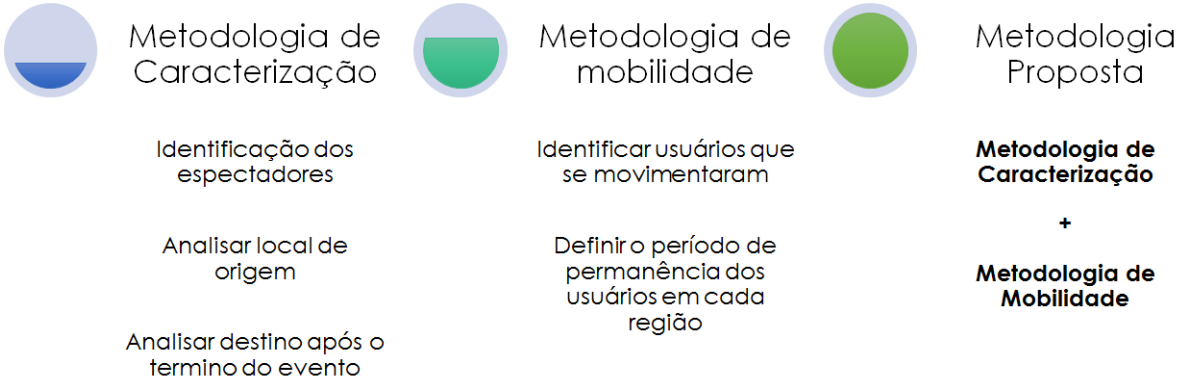


Figura 1 – Metodologia Proposta

4.1 Metodologia de Caracterização

A metodologia de caracterização foi projetada com a finalidade de possibilitar uma visão temporal do uso da rede em grandes eventos. Desta forma, os resultados de sua aplicação podem contribuir para o gerenciamento de uso e planejamento da infraestrutura necessária para grandes eventos. A metodologia utiliza como entrada chamadas de telefones celulares, com usuários anonimizados, mas, unicamente identificados, explicitando as estações rádio bases da operadora (i.e., suas respectivas localizações geográficas) da operadora utilizadas quando do início e término de cada chamada.

A primeira versão da metodologia de caracterização foi desenvolvida especificamente para partidas de futebol (XAVIER et al., 2012). Para isso, criou-se uma forma de analisar a dinâmica da carga de uma rede de telefonia móvel durante esses eventos. A primeira versão estabelecia valores de 45 minutos para os períodos de análise. Este valor foi escolhido por representar o tempo da metade de uma partida de futebol. Entretanto, tais valores não possibilitam analisar eventos de naturezas distintas como shows e festas, onde a duração do período de análise é variável. Desta maneira, criou-se a segunda versão

da metodologia de caracterização. Esta nova versão foi reescrita visando tornar a metodologia genérica, possibilitando assim sua aplicação em eventos de naturezas distintas, tais como shows, festas de Réveillon e corridas de Fórmula 1. Neste trabalho será apresentada a segunda versão da metodologia de caracterização.

4.1.1 Tratamento de Dados

A metodologia de caracterização tem como etapa inicial identificar os usuários de uma operadora de telefonia celular que estiveram envolvidos direta ou indiretamente com o evento de larga escala (ou seja, participantes do evento). Esses usuários serão aqui denominados *participantes*. A premissa principal para definir um usuário como participante consiste na utilização de algum serviço de telefonia móvel, próximo ao local do evento e dentro de um período preestabelecido. A identificação desses participantes permite analisar o volume de ligações geradas pelos mesmos durante toda a coleta, ou seja, antes e depois do evento.

A metodologia de caracterização depende da definição de duas variáveis, a região do evento e uma janela de tempo. A região geográfica do evento se refere não apenas ao local onde o mesmo será realizado mas também à sua região de influência. Um exemplo desta região foi visto nos jogos da Copa do Mundo onde um perímetro de 2.000 metros dos estádios foi isolado segundo (FIFA, 2013). A segunda variável consiste em uma janela de tempo que delimite temporalmente a realização do evento. Essa janela de tempo deve levar em consideração não apenas a duração do evento mas também o período de locomoção para chegar e sair do local. Considerar esses períodos é fundamental para uma análise que permita estudar o evento não apenas quanto ao período de sua realização. Estender a janela de tempo além do período da realização do evento é necessário sobretudo para a análise da mobilidade dos usuários.

As definições da região do evento e da janela de tempo devem levar em consideração a natureza do evento, tamanho do público esperado e local de realização. Uma vez definida a região de interesse, é possível identificar as Estações Rádio Bases (ERBs) que cobrem essa região, cujas cargas poderão ser afetadas pelo público do evento. A definição da janela de tempo possibilita a caracterização do evento. Para tal, adota-se a notação da linha de tempo mostrada na Figura 2. O período marcado como *Evento* corresponde ao período de realização do evento em si. Os demais períodos são explicados a seguir.

O segundo passo da metodologia consiste em identificar os participantes que fizeram pelo menos uma chamada em uma das ERBs, anteriormente selecionadas, em horário próximo ao evento. Note que, de acordo com a natureza do evento, esses participantes tendem a chegar ao local antes de seu início e podem permanecer mesmo após o término do evento. Durante esses períodos, nos quais os participantes aguardam o início do evento ou

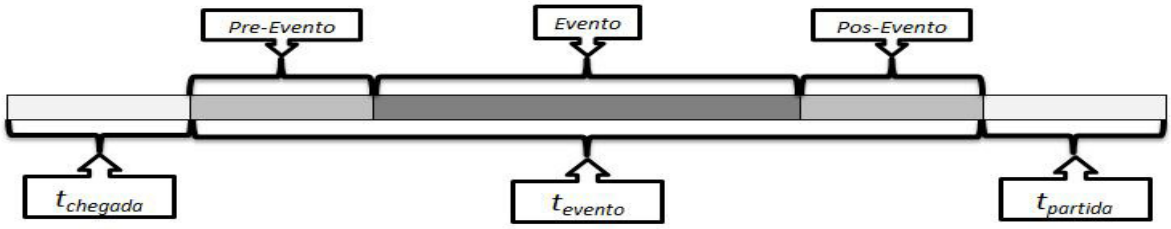


Figura 2 – Linha do tempo para a análise da carga durante um evento

para deixar a região, existe a possibilidade de utilizarem algum recurso da rede de telefonia possibilitando assim sua identificação. Sendo assim, a fim de identificar o maior número possível de usuários envolvidos com o evento foi preciso estender o período que considera a duração do evento analisado. Para tal, é definido o período *Pré-Evento* e o período *Pós-Evento*, que antecede e sucede, respectivamente, o evento. O período total desde o início do *Pré-Evento* até o término do *Pós-Evento* é chamado de t_{evento}. Os usuários que realizaram pelo menos uma chamada durante este período são considerados participantes.

Note que o período *Pré-Evento* deve capturar uma janela de tempo durante a qual os participantes já chegaram no local do evento e estão esperando pelo seu início, enquanto o *Pós-Evento* deve capturar um período durante o qual os participantes ainda estão nos arredores do local, preparando-se para partir. As durações destes períodos dependem da natureza do evento, assim como de características locais e culturais do seu público.

O terceiro e último passo da metodologia consiste em identificar o subconjunto de participantes que também fizeram chamadas antes do início ou após o término do evento (t_{evento}). Essa análise permite rastrear os deslocamentos destes usuários antes e depois do evento. Para identificar esses usuários, são definidos dois outros períodos — $t_{chegada}$ e $t_{partida}$ — durante os quais os usuários estão se movimentando para a região do evento ou partindo dela para outro local. Utilizando a localização geográfica das ERBs onde cada chamada se iniciou e terminou, pode-se determinar de onde esses participantes vieram e para onde foram após o fim do evento. Ou seja, é possível estudar a dinâmica da carga nas ERBs localizadas nas principais rotas de acesso à região do evento. A identificação dessas rotas pode ser explorada para diversos serviços direcionados, tais como para propagandas ou orientações do público.

4.2 Metodologia de Mobilidade

O uso da metodologia de caracterização é fundamental para determinar quais usuários estiveram presentes no local do evento. Entretanto, essa metodologia permite apenas uma visualização superficial da mobilidade desses usuários. Visando solucionar esse problema, foi inserida uma nova etapa à metodologia de caracterização denominada de

metodologia de mobilidade. Essa etapa focou no melhor tratamento dos dados para uma análise mais realista da mobilidade dos participantes de eventos de larga escala. O objetivo desta metodologia proposta é entender a dinâmica da movimentação dos usuários na rede de telefonia celular, a partir das chamadas realizadas. Esta informação pode auxiliar não apenas às empresas de telefonia como também as entidades responsáveis pelo planejamento do evento de larga escala.

Para determinar a movimentação dos participantes de eventos de larga escala, durante o período de análise, são necessárias as informações sobre a ERBs onde esses participantes iniciaram e finalizaram suas chamadas além do posicionamento geográfico (i.e. latitude e longitude) das mesmas. A partir do relacionamento desses dados, se torna possível estudar a mobilidade de todos os usuários que utilizaram duas ou mais ERBs distintas. Desta maneira, a nova etapa da metodologia define uma forma de caracterizar os participantes que se movimentaram e uma forma de definir o tempo de permanência de cada um em cada local.

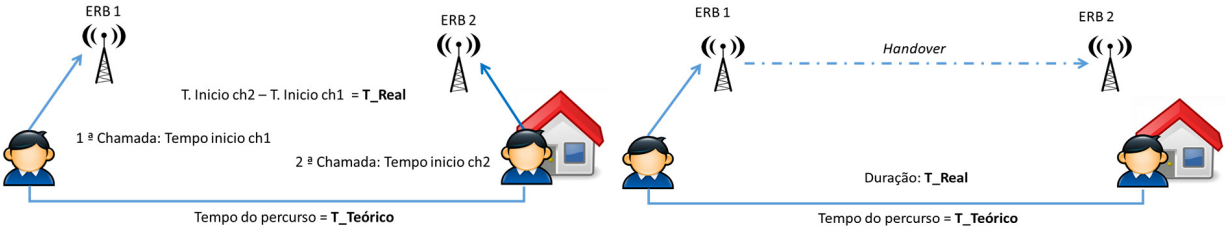
Para a caracterização dos usuários e definição do seu tempo de permanência, inicialmente é definida uma matriz de distâncias. Esta matriz armazena a distância entre cada par de ERBs existentes na base de dados. A partir da matriz de distâncias, são calculados os tempos de permanência dos usuários na região coberta pela ERB. No segundo passo, são selecionadas as chamadas que representam se os usuários movimentaram. Essa caracterização pode ser feita das seguintes maneiras:

1. Identificar a chamada que foi iniciada em uma ERB e finalizada em outra situada em um local diferente ou,
2. Identificar pelo menos duas chamadas do mesmo usuário que iniciaram em ERBs de locais diferentes.

Selecionados os usuários e suas respectivas chamadas, o terceiro passo da metodologia de mobilidade é a definição do tempo que o usuário esteve presente na região. Caso seja possível utilizar apenas uma chamada para identificar a movimentação do usuário, o tempo de permanência na região coberta pela ERB será igual à duração da chamada. Em contrapartida, se para comprovar a movimentação do usuário, forem necessárias duas chamadas, o tempo de permanência será igual ao intervalo entre o início das chamadas. O tempo de permanência, calculado a partir das chamadas do usuário, foi denominado de T_{real}. Os cenários utilizados para o cálculo do T_{real} estão representados na Figura 3.

Uma vez que o tempo de permanência, calculado a partir das chamadas do usuário, na região coberta pela ERB (T_{real}) é conhecido, o próximo passo é calcular o tempo de deslocamento do usuário entre as ERBs utilizadas. Para calcular o tempo de deslocamento dos usuários se faz necessário o conhecimento da velocidade média do deslocamento

em grandes centros. Segundo Munjal, Camp e Navidi (2011) a velocidade média desse deslocamento é de 25 km/h. O tempo médio de deslocamento, entre as ERBs que foram utilizadas durante a chamada do usuário, é calculado pela divisão entre a velocidade média de deslocamento (ou seja, 25 Km/h) e a distância entre as ERBs (armazenado na matriz de distâncias). Esse valor é denominado $T_{previsto}$.



(a) Usuários que realizam chamadas em ERBs distintas **(b) Usuários que se movem enquanto realizam a chamada**

Figura 3 – Cenários de cálculo do T_{real}

Uma vez conhecidos os valores para T_{real} e $T_{previsto}$, o último passo da metodologia de mobilidade consiste em calcular a subtração entre T_{real} pelo $T_{previsto}$. O valor resultante desta subtração representa o tempo em que o usuário permaneceu em cada uma das ERBs utilizadas durante a chamada. Caso o valor resultante dessa subtração seja maior que zero, o tempo de permanência na região coberta pelas ERBs, foi definido empiricamente como como sendo 50% do resultado da subtração na ERB de origem e 50% do resultado da subtração na ERB destino.

As situações onde o resultado dessa subtração seja um valor negativo podem indicar que o usuário estava em uma zona de fronteira. Nesses casos, existe a tendência que o T_{real} possua um valor menor que o $T_{previsto}$. Como a definição do real local de onde o usuário esteve é muito difícil, o T_{real} foi utilizado para determinar o tempo de permanência do usuário nas respectivas regiões. Desta forma, quando o resultado da subtração for um valor negativo, o tempo de permanência é definido como sendo 50% do T_{real} na região ERB de origem e 50% do T_{real} na região da ERB de destino.

A aplicação das metodologias de caracterização e mobilidade nos permite definir as principais vias de acesso e saída, os locais para onde os participantes vão após o evento e quais as regiões de origem desses usuários de telefonia celular. Essas informações são essenciais para um planejamento ideal do acesso e saída do local do evento. Uma vez executadas as metodologias descritas acima, as chamadas realizadas por participantes nas ERBs da região do evento durante t_{evento}, bem como sua mobilidade durante os períodos $t_{chegada}$ e $t_{partida}$ são analisadas, conforme descrito no Capítulo 5.

5 RESULTADOS

5.1 Aplicação da Metodologia nas Bases de Dados

Conforme descrito no Capítulo 4, o primeiro passo da metodologia de caracterização consiste em definir o período de análise e a região de influência do evento. Definido o perímetro de influência do evento, se pode então selecionar as ERBs que realizam sua cobertura. A seleção dessas ERBs é fundamental, pois, durante a análise de eventos na mesma região serão utilizadas as mesmas ERBs, independentemente do tipo de evento realizado. Para todas as bases analisadas (partidas de futebol, Réveillon e shows) foi definida que a região de influência possui um raio no valor de 1.500 metros a partir do local do evento. Note que ao contrário dos jogos da Copa das Confederações e Copa do Mundo onde foi estabelecido um valor de 2.000 metros para a região do evento FIFA (2013), durante a análise das bases de dados será utilizado um valor menor. Essa escolha se deve a análise de eventos de naturezas distintas e jogos de menor importância e representa o raio de análise para estádio Engenhão primeiras bases utilizadas. Sendo assim, para as análises deste trabalho, a região de influência foi reduzida em 500 metros, visando garantir a cobertura de toda a região que sofre alguma influência pela realização de um evento de larga escala.

Quanto a linha de tempo apresentada na Figura 2, o período total adotado para as partidas de futebol foi de 285 minutos. Deste valor, 195 minutos são referentes ao t_{evento} e os outros 90 minutos divididos igualmente entre o $t_{chegada}$ e $t_{partida}$. Como descrito na metodologia o t_{evento} é subdivido em três partes *Pré-Evento*, *Evento* e *Pós-Evento*. Nas partidas de futebol tanto o *Pré-Evento* quanto o *Pós-Evento* possuem a duração de 45 minutos. Este valor de 45 minutos foi utilizado por ser a metade da duração de uma partida de futebol. O *Evento* das partidas de futebol é dividido em três partes sendo a primeira os 45 minutos do primeiro tempo do jogo, a segunda os 15 minutos de intervalo e a terceira e última os 45 minutos do segundo tempo, totalizando assim 105. Este valor de 105 minutos somado aos 45 minutos do *Pré-Evento* e *Pós-Evento* totalizam os 195 minutos referentes ao t_{evento}. Consequentemente, foi adotado o mesmo valor do *Pré-Evento* e *Pós-Evento* para o $t_{chegada}$ e $t_{partida}$, ou seja, 45 minutos. Após a definição desses valores, a linha de tempo para as partidas de futebol pode ser vista na Figura 4.

Nos eventos analisados, referentes às bases de Réveillon e shows, foi utilizada uma linha de tempo com 345 minutos. Os períodos de $t_{chegada}$ e $t_{partida}$, assim como nas partidas de futebol, possuem 45 minutos cada totalizando 90 minutos da linha de tempo. Os

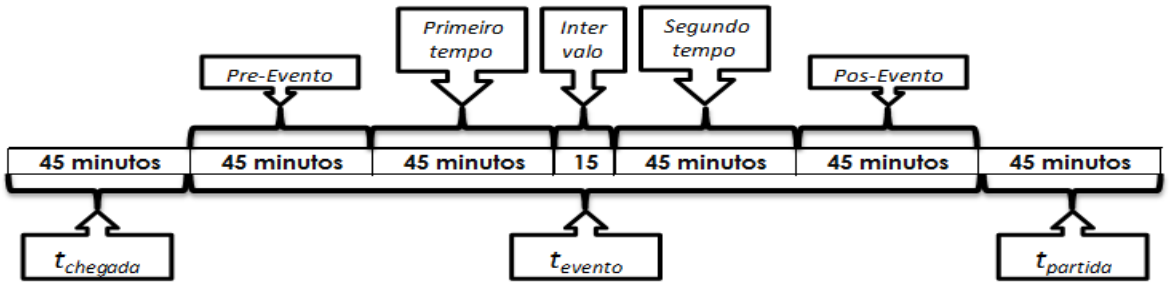


Figura 4 – Linha do tempo para jogos de Futebol

255 minutos restantes são referentes ao t_{evento}. Devido ao grande público esperado para acompanhar os shows e o Réveillon, percebe-se que os participantes cheguem relativamente mais cedo a região do evento. Desta forma, os períodos do *Pré-Evento* e *Pós-Evento* foram acrescidos em 15 minutos, totalizando assim 60 minutos cada. Note que, com a utilização desses valores para os períodos *Pré-Evento* e *Pós-Evento* consideramos que os participantes cheguem com até 1 hora de antecedência ao local do evento e possam permanecer na região pelo mesmo período, após o término do evento. O período *Evento* também foi alterado para as análises de shows e Réveillons. A duração de um show ou da comemoração do Réveillon tende a ser maior que o tempo de uma partida de futebol (105 minutos), para que a linha do tempo contemple o período completo da realizado do evento o intervalo *Evento* foi acrescido em 30 minutos, totalizando assim 135 minutos. Sendo assim, a linha de tempo aplicada as festas de Réveillon e shows é apresentada na Figura 5.

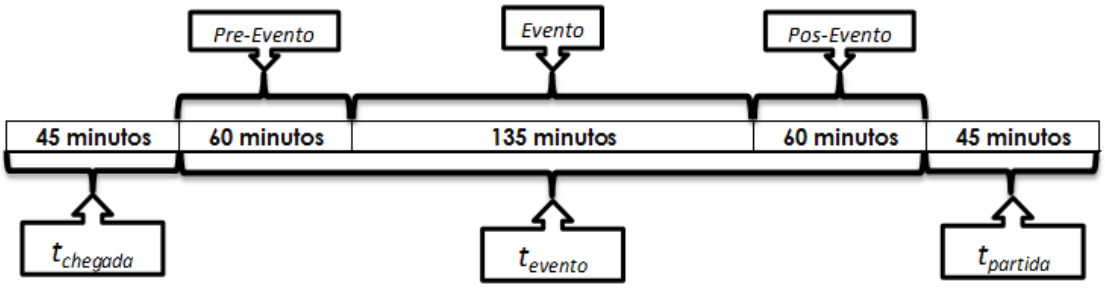


Figura 5 – Linha do tempo para Shows e Réveillon

A duração do GP do Brasil de Fórmula 1 em 2012 foi de 1 hora 45 minutos e 22 segundos, ou seja, 105 minutos e 22 segundos. Esse valor foi muito próximo à duração de uma partida de futebol (105 minutos). Desta maneira, foram adotados os mesmos valores para a linha de tempo de análise do GP do Brasil de Fórmula 1 em 2012. Sendo assim, os períodos de $t_{chegada}$, $t_{partida}$, *Pré-Evento* e *Pós-Evento* possuem 45 minutos e o *Evento* 105 minutos. Entretanto, a corrida não possui divisão durante o *Evento* como nas partidas de futebol. Definidos os valores dos períodos, a linha de tempo para as corridas automobilísticas pode ser vista na Figura 6.

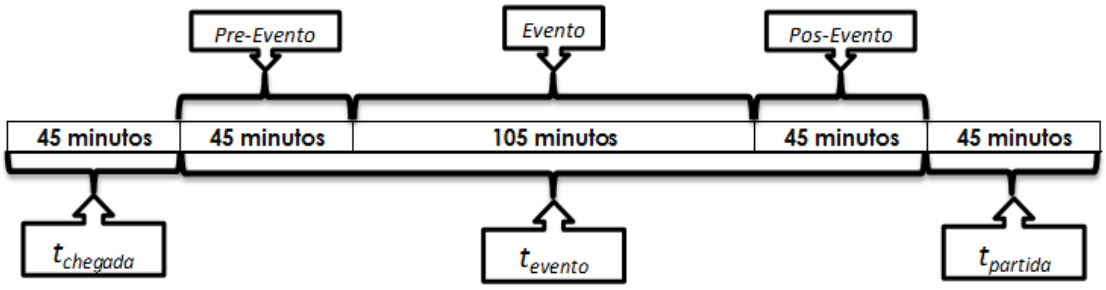


Figura 6 – Linha do tempo para Corridas Automobilísticas

Tendo definido a linha de tempo para cada tipo de evento, a Tabela 5 apresenta os valores utilizados em cada período da linha de tempo aplicados aos respectivos eventos, após execução do primeiro passo da metodologia.

Tabela 5 – Linha de tempo por evento

		$t_{chegada}$		*Pré-Evento*		*Evento*		*Pós-Evento*		$t_{partida}$	
		Inicio	Fim	Inicio	Fim	Inicio	Fim	Inicio	Fim	Inicio	Fim
Réveillon	Belo Horiozonte Recife Rio de Janeiro Salvador São Paulo	21:15	22:00	22:00	23:00	23:00	01:15	01:15	02:15	02:15	03:00
Futebol	Atlético x Fluminense	14:30	15:15	15:15	16:00	16:00	17:45	17:45	18:30	18:30	19:15
	Cruzeiro x Atlético	15:30	16:15	16:15	17:00	17:00	18:45	18:45	19:30	19:30	20:15
	Brasil x Chile	20:30	21:15	21:15	22:00	22:00	23:45	23:45	00:30	00:30	01:15
	Flamengo x Vasco(28)	14:30	15:15	15:15	16:00	16:00	17:45	17:45	18:30	18:30	19:15
	Flamengo x Vasco(04)	15:30	16:15	16:15	17:00	17:00	18:45	18:45	19:30	19:30	20:15
	Flamengo x Fluminense	14:30	15:15	15:15	16:00	16:00	17:45	17:45	18:30	18:30	19:15
	Brasil x Espanha	17:30	18:15	18:15	19:00	19:00	20:45	20:45	21:30	21:30	22:15
	Alemanha x Argentina	14:30	15:15	15:15	16:00	16:00	17:45	17:45	18:30	18:30	19:15
	Holanda x México	11:30	12:15	12:15	13:00	13:00	14:45	14:45	15:30	15:30	16:15
	Grécia x Costa Rica	15:30	16:15	16:15	17:00	17:00	18:45	18:45	19:30	19:30	20:15
	Colombia x Uruguai	15:30	16:15	16:15	17:00	17:00	18:45	18:45	19:30	19:30	20:15
	Flamengo x Atlético PR	20:30	21:15	21:15	22:00	22:00	23:45	23:45	00:30	00:30	01:15
Shows	Beyoncé Paul McCartney Roger Waters	19:45	20:30	20:30	21:30	21:30	23:45	23:45	00:45	00:45	01:30

5.1.1 Visão Geral das Bases de Dados

Com a aplicação da metodologia de caracterização é possível selecionar os participantes e analisar seu comportamento durante cada evento de larga escala. Como descrito no Capítulo 4, participantes são aqueles usuários que realizaram pelo menos uma chamada durante o *Pré-Evento*, *Evento* ou *Pós-Evento*, ou seja, ao longo de todo t_{evento}.

Entender a representatividade desses usuários na base é fundamental para um planejamento eficiente da rede de telefonia móvel. A Tabela 6 mostra o total de usuários e suas respectivas chamadas, além dos mesmos dados referentes aos participantes. As bases de dados referentes aos shows da cantora Beyoncé e do músico Roger Waters apresentaram problemas de limitação da ferramenta de coleta. Durante o show da cantora Beyoncé apenas 72 participantes foram encontrados na região do evento. Tendo apresentado um

público de 20.000 pessoas esse valor não possibilitaria um estudo detalhado dos participantes do evento. O show do músico Roger Waters, mesmo apresentando valores que possibilitariam a caracterização dos usuários, não apresentou nenhuma chamada durante o t_{evento} inviabilizando assim o estudo. Esses problemas serão exibidos detalhadamente nas próximas subseções. Por ser uma tarefa de responsabilidade da operadora, os motivos e propostas de solução para esses problemas não serão tratados neste trabalho.

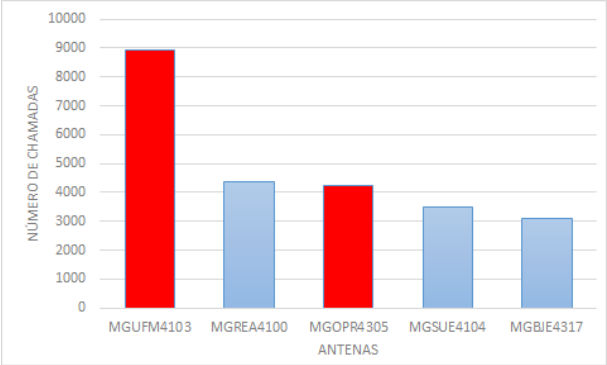
Tabela 6 – Valores totais das bases

		Chamadas			**Usuários**		
		Base	Partic.	% Partic.	Base	Partic.	% Partic.
Réveillon	Belo Horiozonte	19.035	12.800	67%	8.938	5.121	57%
	Recife	24.905	17.016	68%	10.350	5.506	53%
	Rio de Janeiro	38.567	14.707	38%	16.434	4.984	30%
	Salvador	75.285	28.028	37%	31.975	9.617	30%
	Salvador	66.498	33.474	50%	27.189	11.729	43%
Futebol	Atlético x Fluminense	31.705	21.005	66%	13.237	6.868	52%
	Cruzeiro x Atlético	66.963	15.458	23%	33.033	6.339	19%
	Brasil x Chile	45.168	7.276	16%	26.628	3.598	14%
	Flamengo x Vasco(28)	73.711	11.576	16%	38.205	4.625	12%
	Flamengo x Vasco(04)	84.902	14.434	17%	42.144	5.164	12%
	Flamengo x Fluminense	7.579	1.055	14%	4.384	577	13%
	Brasil x Espanha	4.163	1.404	34%	2302	584	25%
	Alemanha x Argentina	10.074	3.468	34%	4.398	1.195	27%
	Holanda x México	8.282	1.012	12%	3.360	340	10%
	Grécia x Costa Rica	13.332	5.709	43%	6.813	2.584	37%
	Colombia x Uruguai	7.305	3.395	46%	3.569	1.434	40%
	Flamengo x Atlético PR	17.007	12.818	75%	7.192	4.839	67%
Shows	Beyoncé	19.024	72	0%	11.620	55	0%
	Paul McCartney	47.931	8.961	19%	25.041	3.083	12%
	Roger Waters	31.166	2.889	9%	18.861	1.459	8%

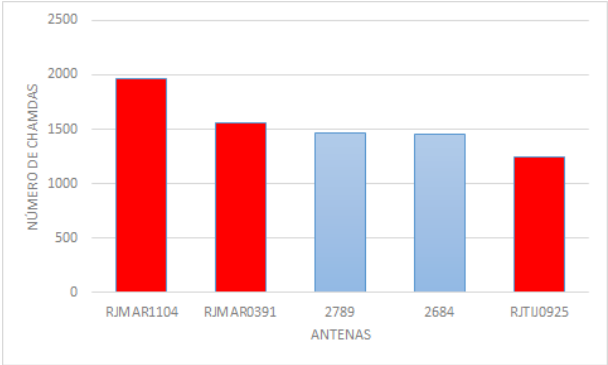
A análise do percentual de usuários que foram considerados participantes das demais bases de dados, apresentado na Tabela 6, mostrou que a metodologia de caracterização identificou que em média 30% dos usuários da base são participantes apresentando um desvio padrão de 17,7. Os dados da Tabela 6 indicam também que a natureza do evento está diretamente ligada ao percentual de participantes coletados. Como a identificação dos participantes depende da existência de pelo menos uma chamada, eventos que motivem os usuários a realizar chamadas, como as festas de Réveillon, apresentam um percentual de participantes maior que os demais.

Os jogos decisivos da Copa do Mundo, Copa das Confederações e Copa do Brasil, apresentam um alto percentual de participantes entre os usuários presentes nas bases de dados. Acreditamos que essas partidas possam ser consideradas como *partidas festivas*. Tais partidas podem atrair participantes de perfis diferentes dos demais jogos. Os resultados nos permite concluir que os participantes desses jogos festivos realizam mais chamadas de seus aparelhos móveis se comparados aos participantes das outras partidas de campeonatos onde os participantes se envolvem muito mais com a partida do que com a socialização decorrente dela.

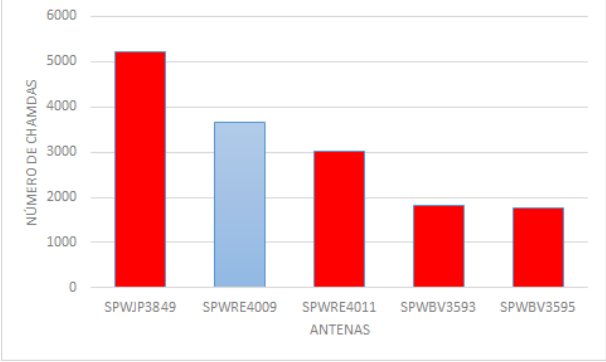
A análise dos dados permitiu identificar um comportamento similar em todas as ERBs que cobrem a região do evento. O levantamento das ERBs mais utilizadas em cada base de dados mostrou que, em geral, pelo menos duas delas fazem a cobertura da região do evento. A Figura 7 apresenta os resultados com as cinco ERBs mais utilizadas na base de dados do Réveillon em São Paulo, na final da Copa do Mundo e na reinauguração do Mineirão. As ERBs em vermelho são aquelas que realizam a cobertura da região do evento. Esses resultados indicam que os participantes possam ser os principais responsáveis pela elevação no número de chamadas da região dificultando assim o planejamento da rede de telefonia móvel. Sabendo deste comportamento,f todos os demais resultados apresentados neste trabalho estão focados nas ERBs que cobrem a região do evento.



(a) Top 5 ERBs - Cruzeiro x Atlético



(b) Top 5 ERBs - Alemanha x Argentina



(c) Top 5 ERBs - Réveillon São Paulo

Figura 7 – Ranking de ERBs mais utilizadas

5.2 Análise das Características do Evento

A versão da metodologia de caracterização descrita no Capítulo 4 permite sua aplicação para definir as características de eventos de natureza distinta. Essa análise estuda o comportamento dos participantes identificados, conforme descrito na Seção 4.1.1,

sendo realizada se concentrando nos seguintes aspectos: número de chamadas, tempo entre início e término de chamadas sucessivas em uma mesma ERB e duração das chamadas. Note que todas as análises consideram apenas os participantes.

A análise dos tempos entre início e término de chamadas sucessivas é de grande valor para investigar como o número de chamadas finalizadas se aproxima ao número de chamadas iniciadas. Essa análise é importante para as operadoras, pois uma sobrecarga em suas ERBs proporciona uma queda na qualidade do serviço prestado. Esses problemas podem refletir em perda financeira além do estabelecimento de uma imagem negativa da empresa. Além disso, a análise da duração das chamadas é importante para entender a ocupação das ERBs durante o período considerado.

5.2.1 Número de Chamadas

Conhecer o volume de chamadas que será atendido pelas ERBs que cobrem a região do evento pode prover importantes informações que auxiliem o planejamento adequado da infraestrutura necessária para suportar eventos de larga escala. Em Tanahashi et al. (2012) os autores indetificaram que o número de chamadas realizadas pelos usuários da rede de telefonia móvel segue uma lei de potência, ou seja, a maior parte dos usuários coletados realizam um número baixo de chamadas. Para a verificação desse mesmo comportamento nas bases de dados analisadas, foi selecionada uma coleta de cada tipo de evento (Show, F1 e Réveillon) além de duas partidas de futebol sendo uma válida pela Copa do Mundo de 2014.

Inicialmente foi calculada a distribuição de probabilidade de cada base de dados como mostrado na Figura 8. Essa distribuição indica a probabilidade de um usuário realizar um número específico de chamadas, durante o período coletado. Em Clauset, Shalizi e Newman (2009) é proposto um método para verificação se uma distribuição segue ou não uma lei de potência. Este foi o método utilizado nas análises das bases de dados e consiste em incialmente levantar a hipótese que todas as distribuições seguem uma lei de potência.

A verificação se a hipótese é verdadeira é realizada a partir da técnica de reamostragem *bootstrap* e o teste de validação de distribuições de probabilidade Kolmogorov-Smirnov, ambos implementados na biblioteca poweRlaw publicada por Gillespie (2014) para o software R-project.

A validação da hipótese foi realizada a partir da propriedade *p-value*, como proposto em Clauset, Shalizi e Newman (2009). Os resultados obtidos a partir do *bootstrap* e do teste de Kolmogorov-Smirnov possibilitam a utilização dessa propriedade. Em Clauset, Shalizi e Newman (2009) é proposto que se utilize um nível de significância (α) de 0,1.

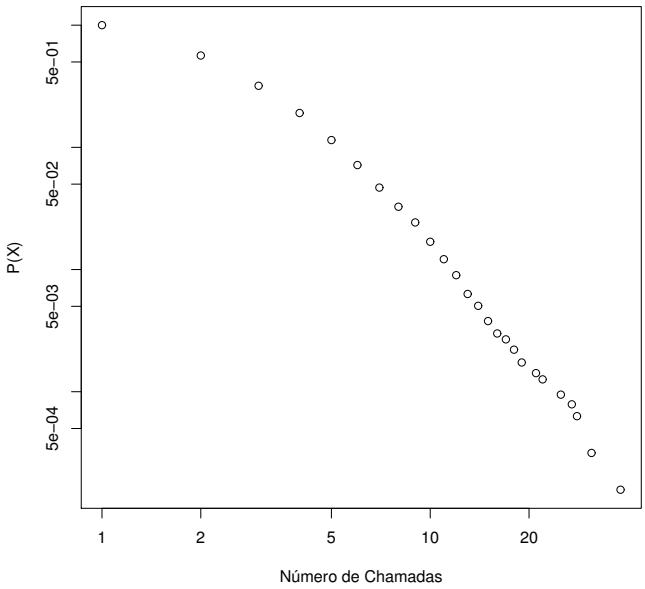


Figura 8 – P(x) número de chamadas - Cruzeiro x Atlético

Desta maneira, caso os resultados obtidos pelos testes apresentem um *p-value* maior ou igual a 0,1 a distribuição segue uma lei de potência. Situação onde os resultados apresentem um *p-value* menor que 0,1, indica que a distribuição não segue uma lei de potência, entretanto, serão necessários novos estudos para analisar qual a melhor distribuição de probabilidade que representa o comportamento da curva.

Como dito anteriormente, foram analisados o Réveillon de São Paulo, o show do músico Paul McCarnety e o GP Brasil de F1 além das partidas de futebol entre Cruzeiro e Atlético entre Grécia e Costa Rica. A Tabela 7 apresenta os resultados das análises dessas bases de dados. Com exceção da partida entre Grécia e Costa rica, os resultados das demais bases analisadas mostrou que o comportamento dos usuários quanto a realização de chamadas segue uma lei de potência, resultado similar ao encontrado em (TANAHASHI et al., 2012).

Tabela 7 – *P-Value* - Número de Chamadas

Bases	***P-Value***
Cruzeiro x Atlético	0,99
GP Brasil de F1	0,94
Grécia x Costa Rica	0,05
Paul McCarnety	0,34
Réveillon SP	0,10

Além do cálculo do *p-value* foram gerados gráficos que permitem a comparação visual dos resultados obtidos com outras distribuições de probabilidade. Para essa comparação foram selecionadas as distribuições de Poison, Exponencial e Log-Normal. A biblioteca de Gillespie (2014) possibilita determinar se uma distribuição discreta segue ou não uma lei de potência. Entretanto, isto só é possível com um valor de x mínimo. Os gráficos da Figura 9 permitem identificar que para valores menores ao x mínimo a distribuição Exponencial apresenta resultados similares ao comportamento dos usuários.

Uma análise agregada dessas chamadas separadas em intervalos de tempo é fundamental para compreender melhor o comportamento da carga de trabalho. Esta análise permite uma visualização de partes do evento onde possa existir um comportamento característico ou um aumento no número de chamadas.

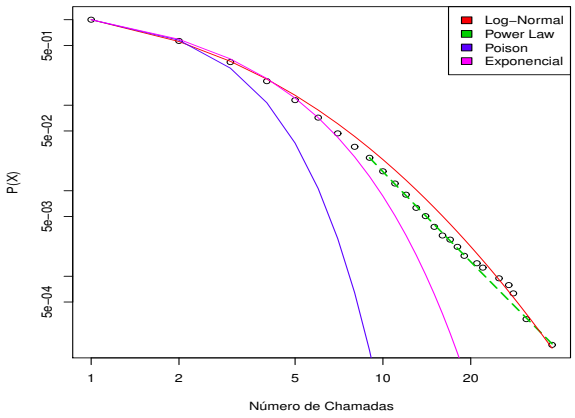
O restante desta seção analisa a distribuição do número de chamadas realizadas pelos participantes. Essas chamadas foram agregadas em intervalos de 15 minutos para melhor visualização durante todo o evento. Os estudos foram realizados durante o período t_{evento} sendo consideradas apenas as ERBs que realizam a cobertura da região do evento. Como mostrado em Bagrow, Wang e Barabási (2011), o tipo de carga de trabalho varia de acordo com a natureza do evento (show, festas, jogos de futebol, etc.). Dessa maneira todas as análises realizadas estão separadas de acordo com a natureza do evento.

5.2.1.1 Partidas de Futebol

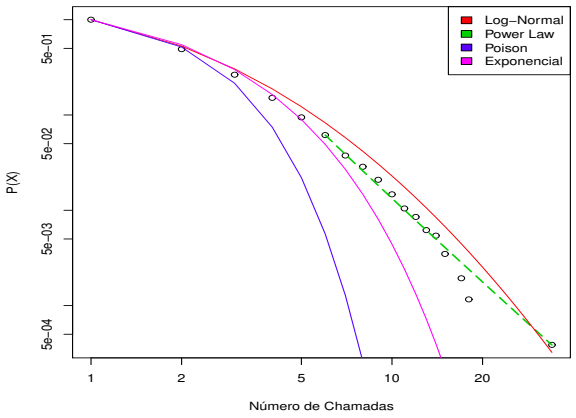
Uma partida de futebol possui um comportamento característico, quando se analisa de forma agregada o número de chamadas realizadas pelos participantes. A Tabela 8 permite identificar três momentos onde os participantes realizam um grande volume de chamadas no decorrer da partida. O primeiro pico ocorre logo no início do *Pré-Evento*, ou seja, 45 minutos antes da partida se iniciar, durante este período os participantes estão realizando chamadas constantemente. Com a proximidade do horário da partida esse volume de chamadas vem reduzindo até sofrer uma forte queda no horário do jogo.

O segundo pico no número de chamadas acontece durante o intervalo da partida. Os 15 minutos que compõem o intervalo (quarta coluna do *Evento*) apresentam o único pico que ocorre durante período *Evento*. O terceiro e último aumento no volume de chamadas acontece logo após o término da partida, início do *Pós-Evento*. Nota-se que, em geral, este pico é maior que os outros dois. Esse comportamento das partidas de futebol foi denominado de efeito "W" devido à similaridade com esta letra, como pode ser observado nas Figuras 10a e 10b.

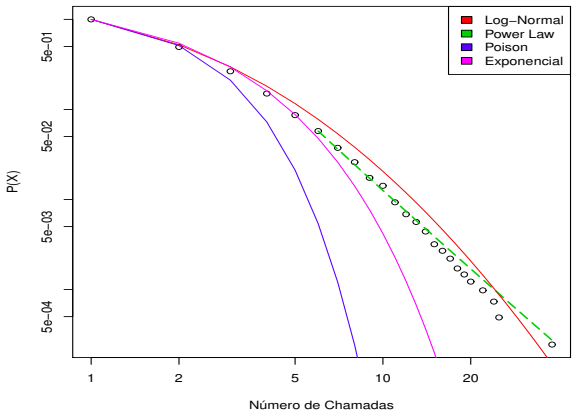
O efeito "W" pode ser notado independente do número de usuários coletados. Durante a partida entre Flamengo e Fluminense do dia 08 de julho, após a aplicação da



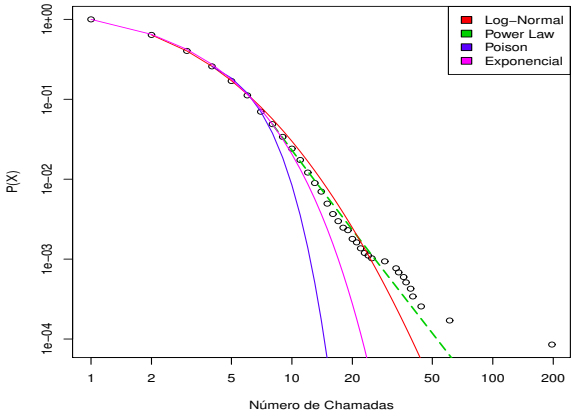
(a) Partida entre Cruzeiro e Atlético



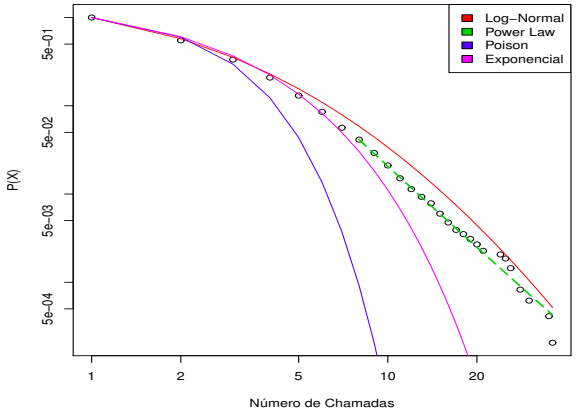
(b) Partida entre Grécia e Costa Rica



(c) Show Paul McCartney



(d) Réveillon São Paulo

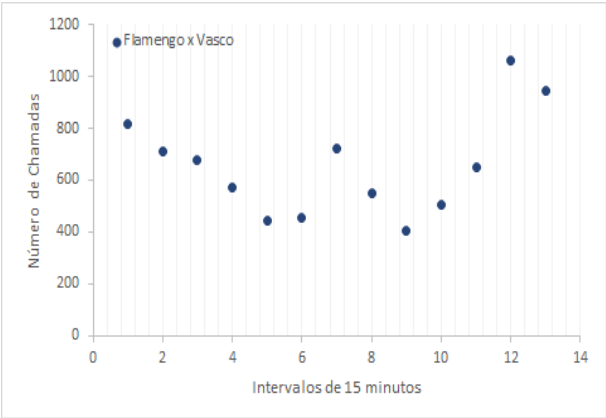


(e) Grande Prêmio do Brasil de F1

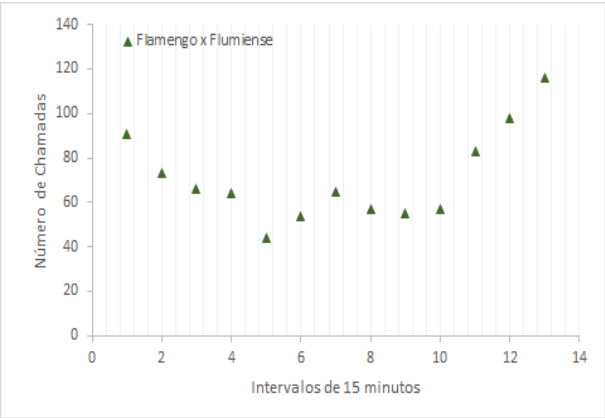
Figura 9 – Gráficos comparativos - Número de chamadas

Tabela 8 – Jogos de futebol – Número de chamadas agregado em intervalos de 15 minutos

	Pré-Evento			*Evento*							*Pós-Evento*		
Bases	***0'***	***15'***	***30'***	***45'***	***60'***	***75'***	***90'***	***105'***	***120'***	***135'***	***150'***	***165'***	***180'***
Cruzeiro x Atlético	53	56	102	782	828	732	1.616	1.273	715	578	1.207	2.143	1.829
Atlético x Fluminense	1.219	1.241	1.226	844	749	632	969	1.091	581	639	905	2.108	1.616
Brasil x Chile	1261	1.191	159	27	24	19	172	847	370	276	377	591	512
Flamengo x Flumiense	91	73	66	64	44	54	65	57	55	57	83	98	116
Flamengo x Vasco	814	711	680	574	441	453	721	548	406	504	649	1.059	942
Flamengo x Vasco (04)	941	997	940	927	892	823	721	562	512	548	748	609	475
Brasil x Espanha	92	93	99	52	39	27	296	106	28	28	91	58	33
Alemanha x Argentina	254	239	209	149	120	112	219	162	112	160	187	134	156
Colombia x Uruguai	268	227	231	204	164	158	211	180	175	126	155	294	171
Grecia x Costa Rica	562	180	525	163	128	558	580	35	493	64	653	231	264
Holanda x México	51	72	66	55	76	56	48	40	34	52	44	38	32
Flamengo x Atlético PR	805	329	206	117	137	333	207	73	49	308	505	-	-



(a) Flamengo x Vasco 28/08



(b) Flamengo x Fluminense

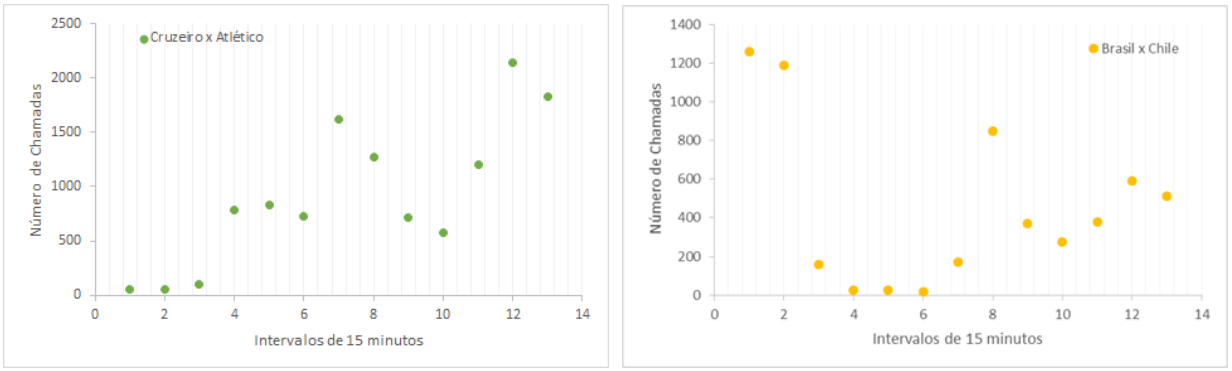
Figura 10 – Efeito "W"

metodologia foram apenas 577 usuários, por outro lado, no jogo entre Flamengo e Vasco do dia 28 de agosto foram encontrados 4.625 participantes. Devido a essa diferença no número de participantes, para melhor visualização dos gráficos apenas o eixo X possui as mesmas escalas. Definir uma escala única para o eixo Y iria inviabilizar a visualização do formato da curva.

As Figuras 10a e 10b mostram que o efeito "W" pode ser visualizado em ambas partidas independente do volume de chamadas geradas. As figuras mostram que a partir do primeiro intervalo de 15 minutos, período em que se inicia o t_{evento}, o primeiro pico no número de chamadas é identificado. O segundo pico, que representa o intervalo da partida, pode ser visualizado no sétimo intervalo de 15 minutos e o terceiro e último pico, após o término da partida, é encontrado no décimo terceiro intervalo e final do *Pós-Evento*.

Conforme dito anteriormente, a natureza do evento está diretamente ligada ao número de chamadas. A análise agregada desta métrica mostrou que o efeito "W" também sofre alterações de acordo com a natureza do evento. Eventos festivos como a reinaugura-

ção do Mineirão e um amistoso da seleção brasileira, apresentaram pequenas variações no formato da curva do número de chamadas. As Figuras 11a e 11b mostram as alterações que a curva "W" sofreu durantes estes dois eventos festivos. Na partida entre Cruzeiro e Atlético podemos perceber que durante o *Pré-Evento* os participantes praticamente não realizaram nenhuma chamada.



(a) Cruzeiro x Atlético **(b) Brasil x Chile**

Figura 11 – Alterações no efeito "W"

O efeito "W" foi caracterizado não apenas nos jogos nacionais. Nas finais da Copa das Confederações de 2013 e da Copa do Mundo de 2014 o mesmo comportamento foi notado. Comportamento apresentado na Figura 12, sobretudo na final da Copa do Mundo onde efeito "W" é mais evidente. A partida de futebol entre Brasil e Espanha o efeito "W" apresentou pequenas variações no formato da curva. Pode-se supor que o envolvimento dos participantes com a partida seja um fator motivador para essas alterações.

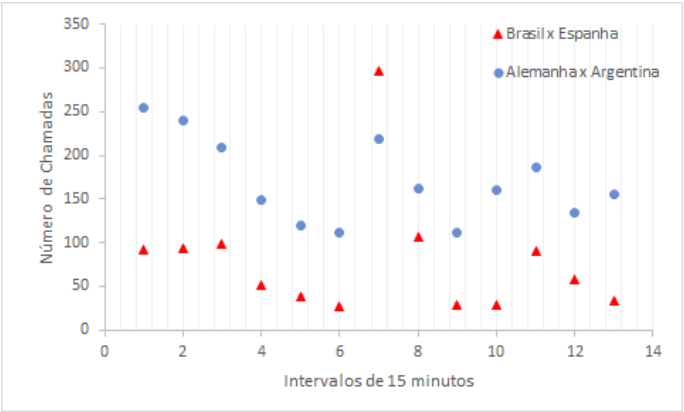
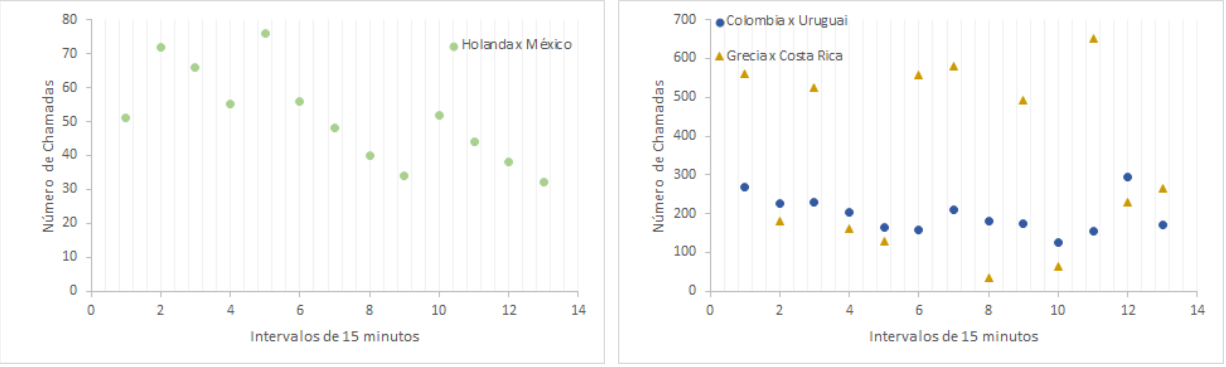


Figura 12 – Número de chamadas - Jogos decisivos no Maracanã

Segundo dados da FIFA foram vendidos 772.232 ingressos para pessoas residentes em países diferentes do Brasil, este valor representou aproximadamente 40% dos ingressos vendidos para os jogos da Copa do Mundo. Desta maneira, a análise das partidas da Copa do Mundo possibilitou o estudo do comportamento de participantes de diferentes

nacionalidades. Dentre os dados analisados, partidas onde nenhuma seleção sul-americana estava presente apresentou um comportamento distinto dos demais. Os resultados das análises destas partidas estão apresentados nas Figuras 13a e 13b. As partidas entre Holanda e México, entre Grécia e Costa Rica e entre Colômbia e Uruguai apresentam comportamento diferente entre si. Pode-se notar que a partida entre Colômbia e Uruguai possui características próximas às outras partidas analisadas.



(a) Holanda x México **(b) Uruguai x Colômbia e Grécia x Costa Rica**

Figura 13 – Efeito "W" nos Jogos da Copa de 2014

5.2.1.2 Shows

Os dados coletados da rede de telefonia móvel durante o show do músico Paul McCartney possibilitaram identificar o comportamento de seus participantes durante o evento. Esses dados podem ser vistos na Tabela 9, juntamente com os dados dos shows do Roger Waters e da Beyoncé. A Figura 14 mostra que os usuários realizam um número maior de chamadas no início e término do evento.

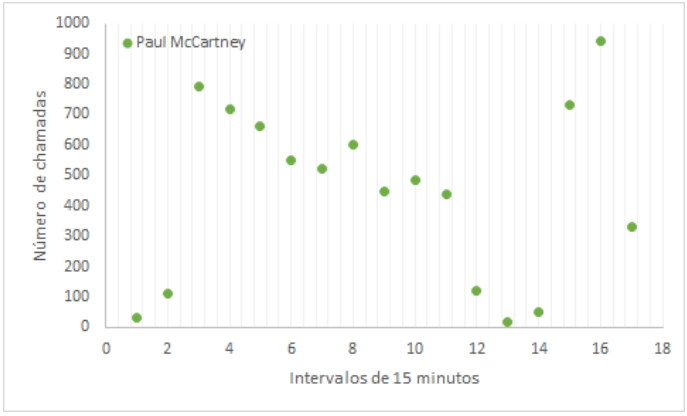
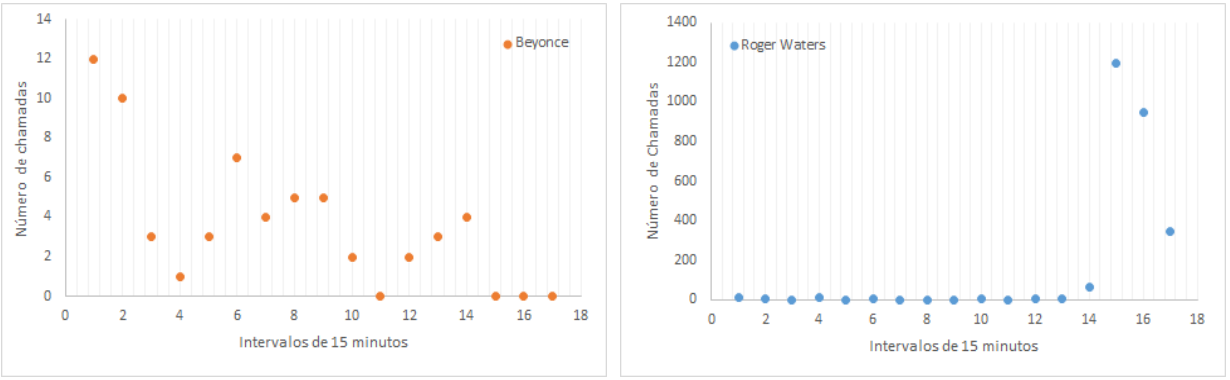


Figura 14 – Número de chamadas - Show Paul McCartney

Tabela 9 – Shows – Número de chamadas agregado em intervalos de 15 minutos

Períodos		Número de chamadas em cada base de dados		
		Roger Waters	Paul McCartney	Beyoncé
Pré-Evento	***0'***	14	31	12
	15'	4	110	10
	30'	1	793	3
	45'	12	717	1
Evento	***60'***	1	663	3
	75'	3	551	7
	90'	1	520	4
	105'	0	600	5
	120'	0	446	5
	135'	2	485	2
	150'	0	438	0
	165'	2	120	2
	180'	2	20	3
Pós-Evento	***195'***	64	49	4
	210'	1198	734	0
	225'	948	941	0
	240'	344	331	0

Após a aplicação da metodologia, os resultados indicaram que os shows da cantora Beyoncé e do cantor Roger Waters apresentaram problemas durante a coleta. As Figuras 15a e 15b apresentam a análise agregada das chamadas dos participantes dessas bases de dados.



(a) Show Beyoncé **(b) Show Roger Waters**

Figura 15 – Número de chamadas shows

Conforme as imagens mostram, após a meia noite, a coleta da rede de telefonia móvel não encontrou nenhuma chamada dos participantes durante o show da cantora Beyoncé. No show do cantor Roger Waters, durante quase todo o t_{evento}, praticamente nenhuma chamada foi coletada. Esses resultados não permitiriam levantar nenhuma hipótese sobre as características do comportamento dos participantes durante esses eventos. Dessa maneira, as bases de dados referentes aos shows da cantora Beyoncé e do cantor Roger Waters foram desconsideradas das próximas análises.

5.2.1.3 Réveillon

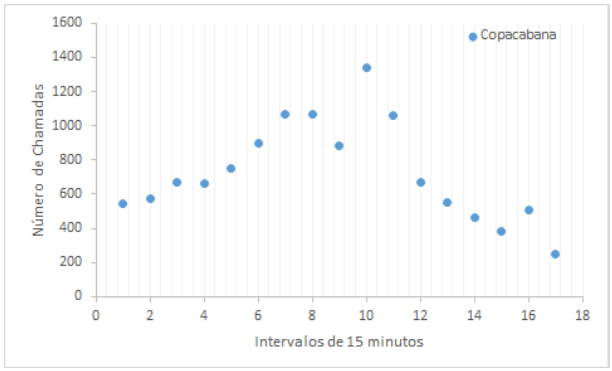
As análises inicias sobre o comportamento dos participantes das festas de Réveillon mostraram que um maior número de chamadas é realizado durante o t_{evento}. Esses dados são apresentados na Tabela 10. Porém, a análise agregada em intervalos de 15 minutos, permite perceber um comportamento distinto nos intervalos próximos à meia noite. As análises mostram que nesse período, no período que o antecede ou no período subsequente à meia noite, acontece uma redução no número de chamadas.

Tabela 10 – Réveillon – Número de chamadas agregado em intervalos de 15

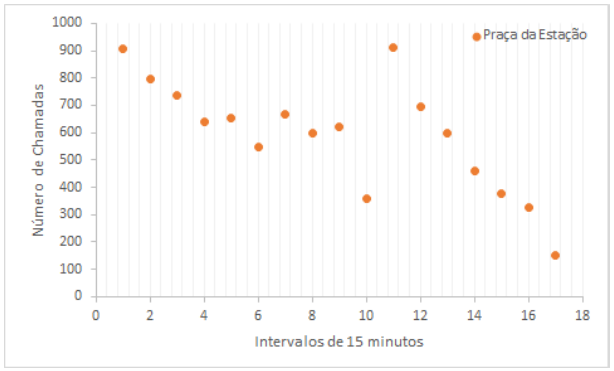
		Número de chamadas em cada base de dados				
Períodos		Copacabana	Praça da Estação	Salvador	Recife	São Paulo
Pré-Evento	***0'***	542	909	1591	942	1815
	15'	575	797	1678	871	1680
	30'	671	736	1539	835	1674
	45'	665	639	176	847	1003
Evento	***60'***	753	655	1898	802	1641
	75'	900	547	1904	797	1613
	90'	1069	667	893	750	1553
	105'	1068	597	135	806	1795
	120'	882	623	724	964	1732
	135'	1342	358	1478	858	678
	150'	1059	910	1550	727	2705
	165'	672	696	1191	611	1853
	180'	554	598	1448	576	1879
Pós-Evento	***195'***	466	462	1367	542	1349
	210'	385	379	1183	467	1295
	225'	510	327	1181	489	1073
	240'	248	152	1215	422	676

Esse comportamento pode ser observado nas Figuras 16a, 16b e 16c. Na Figura 16a se nota que próximo à meia noite, ou seja, nos intervalos oito, nove ou dez do gráfico há uma variação no número de chamadas. Esse comportamento é observado nas análises do Réveillon Belo Horizonte e Salvador. Em Belo Horizonte, o intervalo referente às 00:15 (décimo intervalo) aponta uma grande redução no número de chamadas. Na festa de Réveillon de Salvador, esse comportamento é encontrado às 23:45, ou seja, no oitavo intervalo.

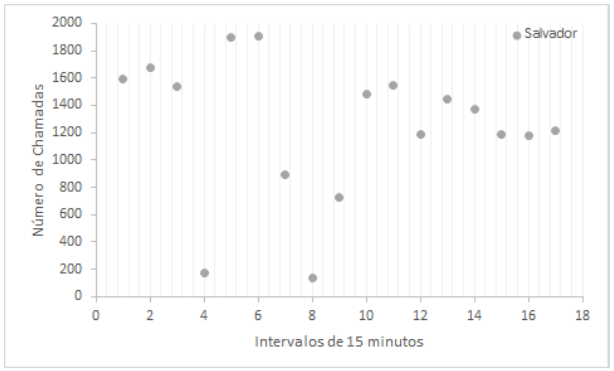
Em eventos de larga escala onde exista um ponto que motive os usuários a realizarem chamadas, como a virada do ano, é perceptível a dificuldade de utilizar os serviços de telefonia móvel. Essa dificuldade pode explicar o comportamento distinto da análise do Réveillon de Salvador. Em diversos pontos da Figura 16c se percebe uma grande redução no número de chamadas, se comparado com o intervalo anterior. Esse comportamento pode indicar que a rede de telefonia móvel do Réveillon de Salvador apresentou limitações para atender aproximadamente 1 milhão de pessoas. As Figuras 16a e 16c permitem visualizar melhor essas possíveis limitações. Note que a curva gerada pelos dados do Réveillon de Copacabana apresenta um comportamento mais constante.



(a) Réveillon Rio de Janeiro



(b) Réveillon Belo Horizonte



(c) Réveillon Salvador

Figura 16 – Número de chamadas Réveillons

Devido ao número reduzido de participantes, a análise de forma agregada do número de chamadas no Réveillon de Recife permite uma melhor visualização do comportamento dos participantes. A Figura 17 mostra esses dados. Percebe-se que o volume das chamadas apresenta um comportamento mais uniforme. Como esperado, o gráfico apresenta um pico a meia noite e ao contrário dos demais Réveillons apresentados na Figura 16, não se pode perceber reduções agudas e sim de forma gradual com o decorrer do tempo.

5.2.1.4 Automobilismo

O último tipo de evento de larga escala analisado neste trabalho foi o Grande Prêmio do Brasil de Fórmula 1 de 2012. A característica do evento se mostrou determinante para a forma como os participantes utilizaram os recursos da rede de telefonia móvel. Durante uma corrida de Fórmula 1 não existem pausas, como um intervalo de uma partida de futebol ou momentos que incentivem os participantes a realizarem ligações como a meia noite de um Réveillon. Sendo assim, a Tabela 11 mostra o comportamento uniforme dos participantes desse tipo de evento. A Figura 18 mostra o comportamento durante o

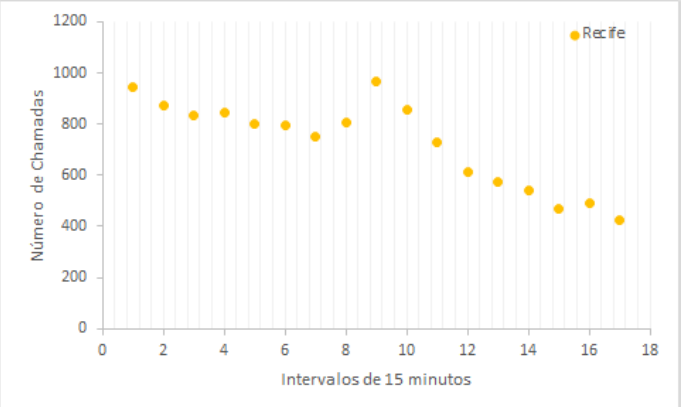


Figura 17 – Número de chamadas Réveillon Recife

Grande Prêmio do Brasil de Fómula 1. Como se pode notar o número de ligações é uniforme durante todo o período analisado ao contrário da partida entre Flamengo e Atlético Paranaense onde podemos notar o efeito "W".

Tabela 11 – Tabela Automobilismo - Análise agregada

	Pré-Evento			*Evento*							*Pós-Evento*		
Bases	***0'***	***15'***	***30'***	***45'***	***60'***	***75'***	***90'***	***105'***	***120'***	***135'***	***150'***	***165'***	***180'***
Formula 1	791	949	764	984	769	658	646	620	568	567	665	689	633

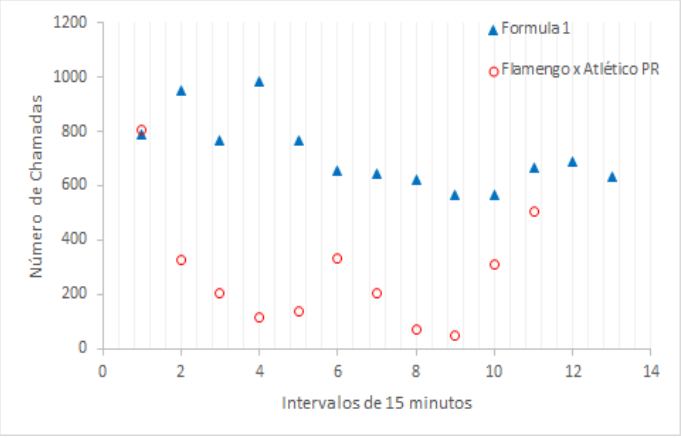


Figura 18 – Número de chamadas GP do Brasil e Final Copa do Brasil

5.2.2 Intervalo entre Chamadas

O tempo entre inícios e o tempo entre términos de chamadas sucessivas em uma mesma ERB, aqui denominados de *InterArrival Times* (IAT) e *InterDeparture Times* (IDT), são fundamentais para avaliar a carga imposta nas ERBs. A partir dessas métricas

é possível analisar as alterações no volume de chamadas iniciadas ou finalizadas no decorrer do tempo. Tais métricas também são essenciais para a criação de cenários de simulação. Os valores encontrados em eventos de naturezas distintas podem auxiliar a simulação do mesmo evento em locais diferentes, como uma final de Copa do Mundo no Mineirão. Todos os valores calculados serão apresentados em três grupos diferentes, partidas de futebol, festas ou show e corrida automobilística.

5.2.2.1 Partidas de Futebol

Todos os valores apresentados na Tabela 12 foram calculados durante o período do t_{evento}, esta Tabela apresenta os valores médios de IAT e IDT. As colunas S_{inter} e S_{intra} representam, respectivamente, o desvio padrão do tempo médio computado para diferentes ERBs e a média dos desvios padrões computados para cada ERB durante todo o período do t_{evento}.

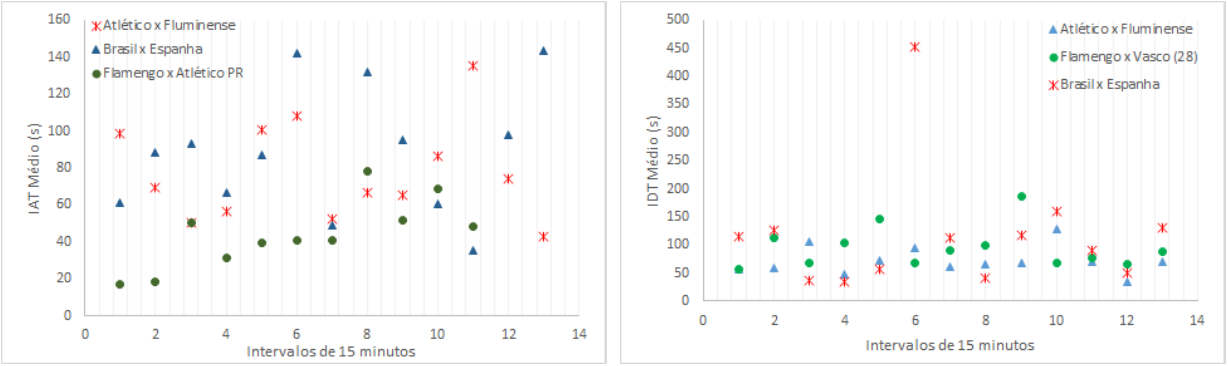
Tabela 12 – Tabela IAT e IDT - Jogos de futebol

		IAT			**IDT**		
		IAT médio (s)	S_{inter}	S_{intra}	IDT médio (s)	S_{inter}	S_{intra}
Futebol	Atlético x Fluminense	100,60	102,61	60,64	94,75	103,80	44,35
	Cruzeiro x Atlético	33,46	25,15	46,20	35,42	15,64	33,80
	Brasil x Chile	37,018	31,41	24,46	64,69	35,23	55,59
	Flamengo x Vasco(28)	112,94	116,69	72,44	121,26	128,52	88,63
	Flamengo x Vasco(04)	97,12	98,68	57,66	114,44	134,21	53,38
	Flamengo x Fluminense	186,95	123,98	124,48	184,47	112,53	103,10
	Brasil x Espanha	105,38	45,27	43,34	170,96	90,15	145,61
	Alemanha x Argentina	121,46	103,27	104,06	253,65	186,64	129,76
	Holanda x México	101,01	41,58	36,97	177,83	90,23	87,06
	Grécia x Costa Rica	32,47	16,22	47,09	71,49	57,87	65,89
	Colombia x Uruguai	122,74	157,45	27,77	226,50	208,39	83,21
	Flamengo x Atlético PR	69,71	76,27	30,15	95,78	99,39	61,16

Assim como na análise do número de chamadas, os resultados indicam que a motivação da partida influencia na característica de uso dos recursos de telefonia móvel. Os dados mostram que, nas partidas festivas entre Cruzeiro e Atlético ou entre Brasil e Chile, as chamadas aconteceram com maior frequência ao longo da partida. Estes eventos apresentaram os menores valores de IAT médio 33,46 e 37,01 segundos respectivamente. Jogos finais, ou considerados assim, como Atlético contra Fluminense, Brasil e Espanha além de Flamengo e Vasco de 04 de dezembro, apresentaram valores maiores no IAT médio. Esse comportamento mostra que durante o período do t_{evento}, o número de chamadas iniciadas reduziu bastante, possivelmente pelo envolvimento do participante com o evento, aumentando assim seu IAT médio.

A Figura 19 mostra a análise agregada, em intervalos de 15 minutos, dos valores médios de IAT e IDT durante as partidas de futebol consideradas como finais. Como se pode perceber na Figura 19a durante o intervalo do jogo os valores de IAT tendem

a diminuir. Esse comportamento indica um maior número de chamadas neste período. A Figura 19b mostra que durante o intervalo os participantes da final entre Brasil e Espanha tiveram um alto valor de IDT. Essas características das chamadas correspondem às chamadas de grande duração e consequentemente maior ocupação dos recursos da rede.



(a) IAT - Final Copa das Confederações, Copa do Brasil e Atlético e Fluminense

(b) IDT - Final Copa das Confederações, Copa do Brasil e Atlético e Fluminense

Figura 19 – IAT e IDT Finais Futebol

Como pode ser visto na Figura 20, em praticamente todos os intervalos de 15 minutos os valores médios de IAT são menores se comparados a partida de futebol entre Flamengo e Vasco referentes ao Campeonato Brasileiro. Esses resultados podem indicar que os participantes dos jogos da Copa utilizaram a rede de telefonia móvel de uma maneira mais constante ao longo de todo o t_{evento}.

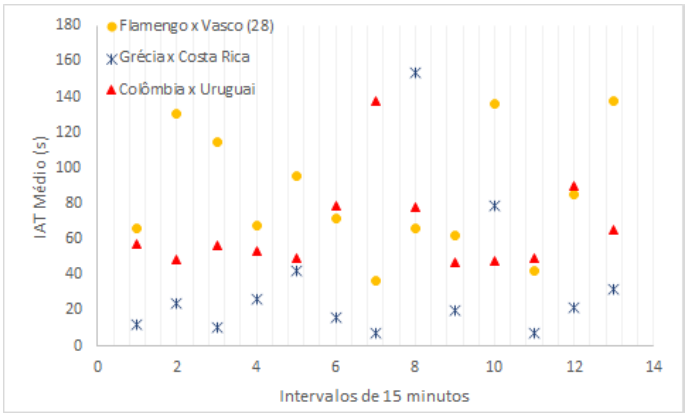


Figura 20 – Comparação do IAT dos jogos da Copa com o Jogo entre Flamengo e Vasco (28)

5.2.2.2 Réveillon e Shows

Assim como na análise das partidas de futebol os valores apresentados na Tabela 13 foram calculados durante o t_{evento}, que mostra os valores médios de IAT e IDT. As colunas S_{inter} e S_{intra} representam, respectivamente, o desvio padrão do tempo médio e a média dos desvios padrões computados para cada ERB, também durante o período do t_{evento}.

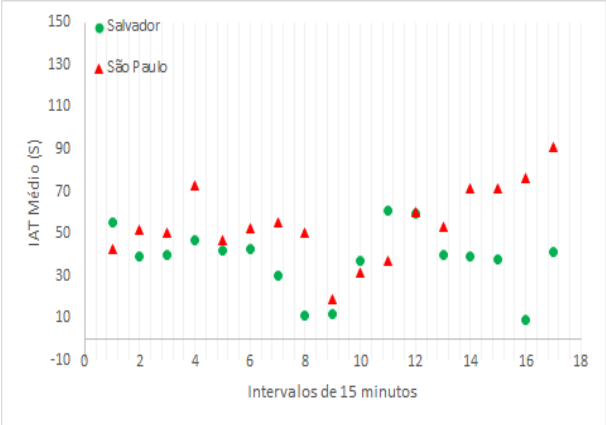
Tabela 13 – Tabela IAT e IDT - Réveillons e show

		IAT			IDT		
		IAT médio (s)	S_{inter}	$\overline{S_{intra}}$	IDT médio (s)	S_{inter}	$\overline{S_{intra}}$
Réveillon	Belo Horiozonte	112,38	21,24	25,58	57,03	22,28	58,53
	Recife	54,06	74,16	26,23	68,39	72,59	51,51
	Rio de Janeiro	114,46	95,48	80,12	235,73	158,35	146,88
	Salvador	50,76	70,42	31,00	66,21	104,87	45,04
	São Paulo	57,39	38,14	29,20	68,66	89,39	30,13
Shows	Paul McCartney	33,42	21,24	25,58	57,03	22,28	58,53

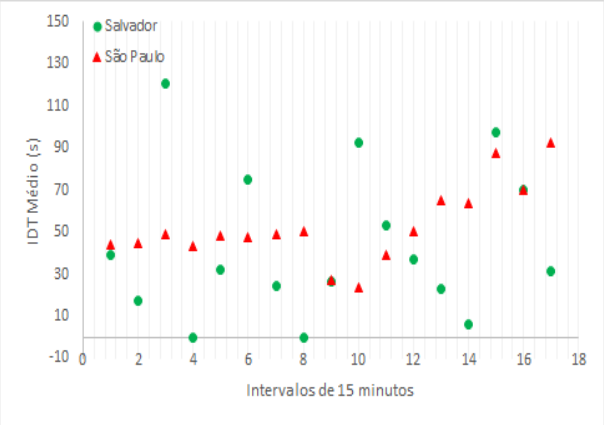
Nota-se que com exceção do Réveillon de Belo Horizonte os valores de desvio padrão são próximos ou maiores que a média de todos os demais eventos analisados. Por serem eventos em locais cobertos por um número maior de ERBs, algumas áreas da região do evento podem estar mais ocupadas que outras gerando assim essa disparidade nos valores médios.

O estudo dos intervalos entre chamadas nos permite identificar também a duração das ligações que são realizadas. Como pode ser analisado nas Figuras 21 nas festas de Réveillon de São Paulo e Salvador os valores de IAT e IDT são próximos. Esse comportamento indica que a ligações realizadas pelos participantes deste evento são de curta duração.

Durante a análise do número de chamadas do Réveillon os resultados indicaram uma possível sobrecarga da rede de telefonia móvel durante a festa de Réveillon em Salvador. A Figura 22 mostra de forma comparativa o número de chamadas do Réveillon de Salvador e o IAT do mesmo evento. Pode-se notar que durante o intervalo iniciado as 23:45, oitavo intervalo, há uma redução no número de chamadas e também no IAT. A redução do IAT deveria refletir em um aumento no número de chamadas, entretanto, o alto número de chamadas iniciadas durante um curto espaço de tempo, pode ter atingindo o limite da rede.

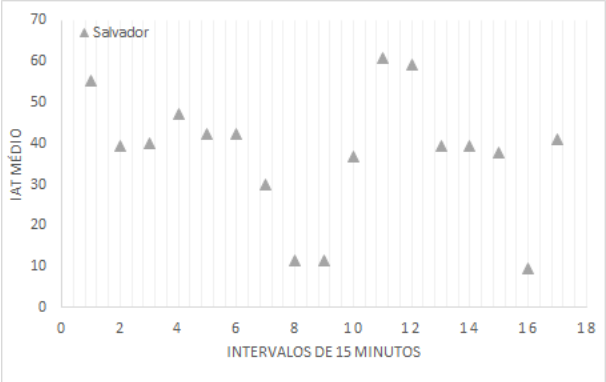


(a) IAT Réveillon São Paulo e Salvador

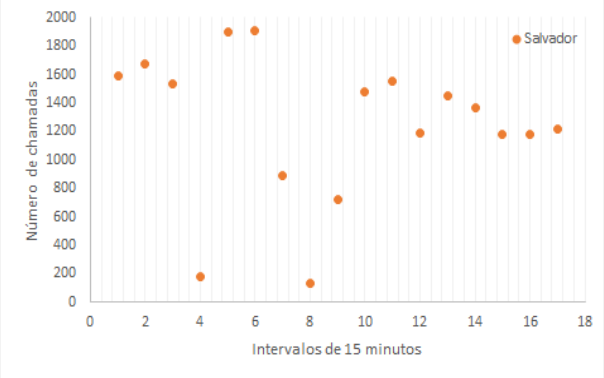


(b) IDT Réveillon São Paulo e Salvador

Figura 21 – IAT e IDT festas de Réveillon



(a) IAT Réveillon Salvador



(b) Chamadas Réveillon Salvador

Figura 22 – Comparativo Número de chamadas e IAT do Réveillon de Salvador

5.2.2.3 Automobilismo

A análise do número de chamadas durante eventos automobilísticos mostrou que mesmo não apresentando nenhum momento que incentive os usuários a realizarem chamadas, os valores encontrados durante o GP do Brasil de 2012 foram menores que uma final de campeonato. A Tabela 14 mostra que além do grande volume os participantes realizaram chamadas constantemente.

A análise agregada mostra que a infraestrutura necessária para receber esse tipo de evento deve ser previamente planejada. A Figura 23 apresenta de forma comparativa os valores de IAT do Réveillon de São Paulo, da final da Copa do Mundo e do GP do Brasil. A análise gráfica permite identificar que os valores de IAT encontrados durante o GP do

Tabela 14 – Tabela IAT e IDT - Automobilismo

	IAT			IDT		
	IAT médio (s)	S_{inter}	S_{intra}	IDT médio (s)	S_{inter}	S_{intra}
GP do Brasil 2012	25,61	40,74	7,03	24,70	38,43	8,77

Brasil são menores aos encontrados durante o Réveillon de São Paulo. Apenas durante o nono período (meia noite), onde o IAT foi de apenas 19,06 segundos, o Réveillon de São Paulo apresentou um valor médio menor. Nota-se que, mesmo em uma quantidade menor, os participantes do GP do Brasil em geral realizam suas chamadas com uma frequência próxima à meia noite de um Réveillon de 2 milhões de pessoas.

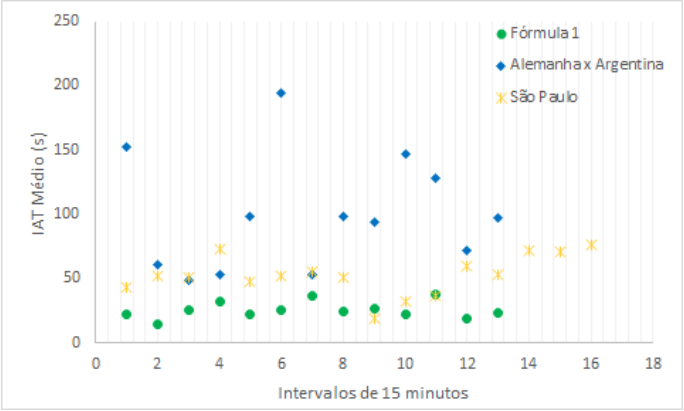


Figura 23 – Comparativo IAT Automobilismo, Réveillon e Partida de Futebol

5.2.3 Duração das Chamadas

A fim de estender os resultados obtidos em Tanahashi et al. (2012), foi analisado tanto a distribuição das chamadas dos participantes quanto a duração das mesmas. Sendo assim, a análise da duração utiliza o mesmo processo para verificação se uma distribuição segue, ou não, uma lei de potência apresentado anteriormente.

A Tabela 15 mostra os resultados dos cálculos do *p-value* em cada base analisada. Os resultados mostram que durante os eventos não esportivos analisados como shows e o Réveillon, não apenas o número de chamadas, mas também a duração seguem uma lei de potência.

Os gráficos apresentados na Figura 24 permitem indentificar que quando se analisa a duração das chamadas, realizadas pelos espectadores, a melhor distribuição de probabilidade que expressa o comportamento dos espectadores quanto a duração de suas chamadas é a Log-Normal.

Tabela 15 – *P-Value* - Duração das Chamadas

Bases	***P-Value***
Cruzeiro x Atlético	0,001
F1	0,001
Grécia x Costa Rica	0,012
Paul McCarnety	0,138
Réveillon SP	0,229

Além dos valores de IAT e IDT, a análise das durações das chamadas dos espectadores é fundamental para a compreensão do uso das ERBs. Como na análise do IAT e IDT os cálculos foram feitos durante períodos de 15 minutos. Chamadas onde a duração ultrapassa o período de 15 minutos são consideradas apenas no intervalo onde a mesma foi criada.

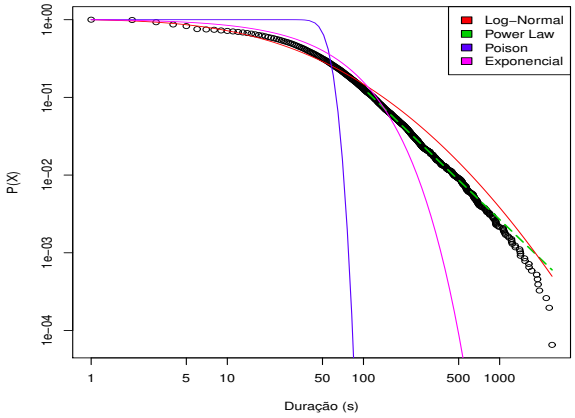
5.2.3.1 Partidas de Futebol

A Tabela 16 apresenta os resultados dos cálculos da duração nos eventos esportivos. Como na análise do IAT e IDT os jogos festivos apresentaram valores que indicam uma maior utilização dos recursos da rede.

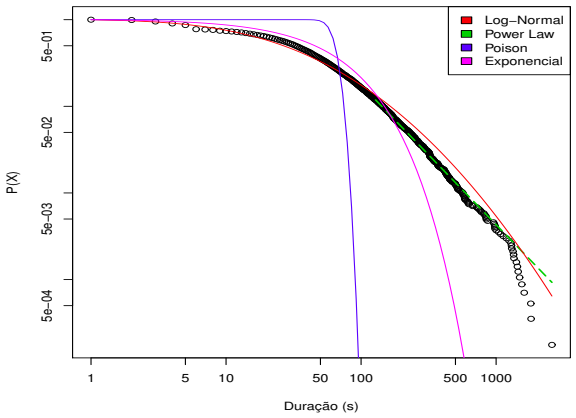
Tabela 16 – Tabela de duração - Jogos de futebol

		Duração		
		Duração média (s)	S_{inter}	$\overline{S}_{intra}$
Futebol	Atlético x Fluminense	82,92	32,16	63,18
	Cruzeiro x Atlético	96,29	67,51	128,10
	Brasil x Chile	227,80	158,81	440,27
	Flamengo x Vasco(28)	70,08	14,54	35,80
	Flamengo x Vasco(04)	74,27	15,48	46,59
	Flamengo x Fluminense	38,17	12,73	30,53
	Brasil x Espanha	67,29	25,87	33,89
	Alemanha x Argentina	153,25	211,66	204,95
	Holanda x México	299,32	347,86	64,21
	Grécia x Costa Rica	92,40	38,01	91,31
	Colômbia x Uruguai	265,71	520,75	343,80
	Flamengo x Atlético PR	69,36	43,22	35,98

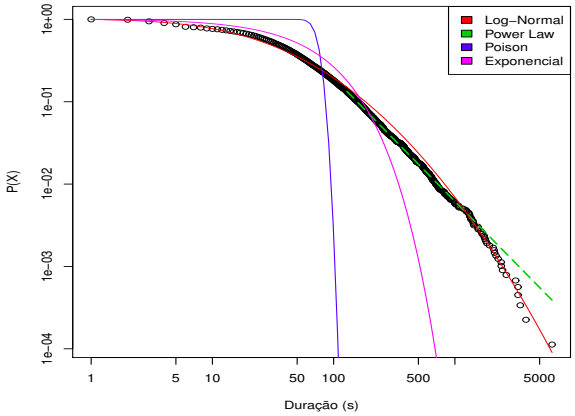
Os participantes do amistoso entre Brasil e Chile realizaram chamadas com uma duração média de aproximadamente 4 minutos. Essa característica, juntamente com os resultados de IAT, que mostrou um grande número de chamadas realizadas em um curto espaço de tempo, faz com que partidas festivas como essa necessitem de um melhor planejamento da capacidade da rede.



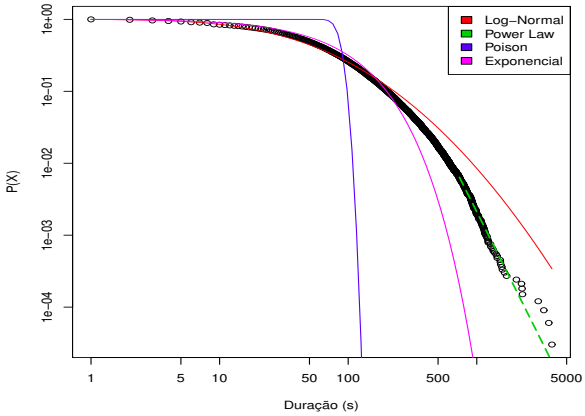
(a) Partida entre Cruzeiro e Atlético



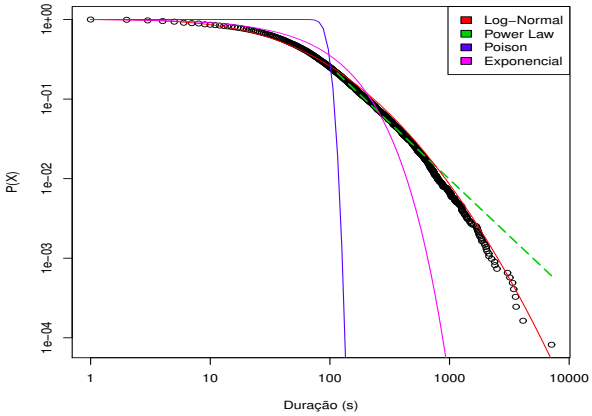
(b) Partida entre Grécia e Costa Rica



(c) Show Paul McCartney



(d) Réveillon São Paulo



(e) Grande Prêmio do Brasil de F1

Figura 24 – Gráficos comparativos - Duração

O comportamento distinto dos participantes presentes na Copa do Mundo pode ser notado também quando se analisa a duração de suas chamadas realizadas. A Figura 25 apresenta um cenário que em praticamente todos os períodos a final da Copa do Mundo, as chamadas dos participantes apresentaram valores mais altos quanto a duração. O período que contempla o intervalo em especial apresentou ligações com duração média de oito minutos. Note que durante o intervalo da final da Copa do Mundo de 2014 foi iniciada uma chamada de grande duração refletindo assim no ponto fora da curva.

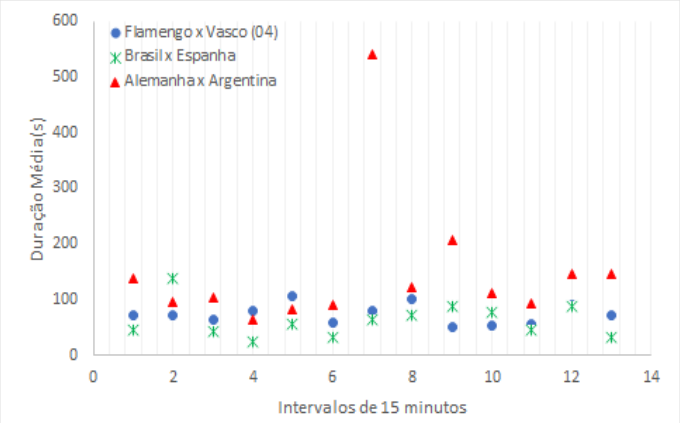


Figura 25 – Duração Chamadas - Finais Maracanã e Flamengo x Vasco (04)

A Figura 26 mostra um comparativo entre duas partidas consideradas como finais. Pode-se notar a similaridade nos valores e no formato da curva. Ao contrário do que acontece com as partidas festivas, o aumento da duração das chamadas nestas partidas acontece nos períodos onde há uma redução no número de chamadas. Desta maneira a rede de telefonia móvel não tende a ficar sobrecarregada, garantindo assim a disponibilidade do serviço aos usuários.

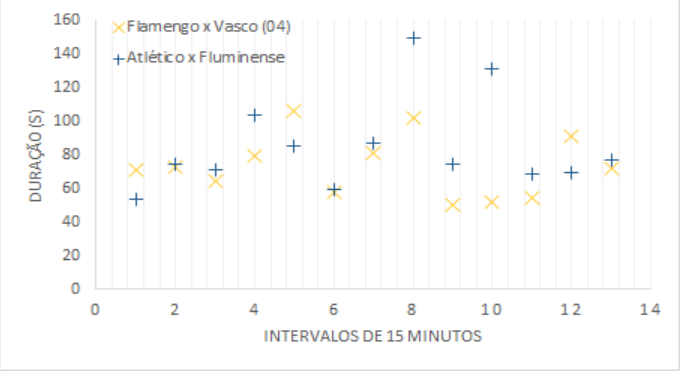


Figura 26 – Duração jogos de futebol

5.2.3.2 Réveillon e Shows

Os valores calculados para os eventos de larga escala analisados são apresentados na Tabela 17. Nota-se um alto valor no desvio padrão durante a festa de Réveillon no Rio de Janeiro. Esse comportamento pode refletir a dificuldade de caracterização do comportamento de um público específico, já que 2,5 milhões de pessoas estiveram presentes no evento. Em comparação com as demais festas de Réveillon o show do Paul McCartney apresentou valores maiores quanto à duração das chamadas, com aproximadamente 2 minutos.

Tabela 17 – Tabela de duração - Réveillons e show

		Duração		
		Duração média (s)	S_{inter}	S_{intra}
Réveillon	Belo Horiozonte	65,30	22,76	40,35
	Recife	58,24	18,94	22,65
	Rio de Janeiro	95,54	72,82	61,67
	Salvador	67,39	27,91	35,89
	São Paulo	98,03	29,41	40,84
Shows	Paul McCartney	114,80	51,46	141,85

A Figura 27 mostra a análise agregada da duração dos eventos. Como podemos notar o Réveillon de Recife apresentou um padrão mais linear ao contrário das festas de Réveillon em Belo Horizonte e Rio de Janeiro. Esse comportamento mostra a dificuldade de caracterizar um comportamento específico para os participantes nesses eventos, onde um grande número de pessoas estarão presentes, e consequentemente planejar a infraestrutura necessária para suportá-lo.

Como mostrado na análise do IAT e IDT os valores próximos nas festas de Réveillon indicaram que as chamadas possuíam curta duração. Entretanto, a Figura 28 mostra que a duração das chamadas durante grande parte do Réveillon é maior se comparada com uma partida de futebol. Contudo, durante o intervalo, sétimo período, a duração das chamadas dos participantes se aproximou dos valores encontrados no Réveillon.

5.2.3.3 Automobilismo

A análise do IAT e IDT durante o GP do Brasil mostrou que eventos dessa natureza apresentaram os menores valores, se comparados aos demais tipos de eventos analisados. Desta maneira, a análise da duração é essencial para uma caracterização destes usuários. A Tabela 18 apresenta os valores médios da duração das chamadas destes participantes.

A Figura 29 permite a análise agregada da duração das chamadas do GP do Brasil. Percebe-se que além de possuir baixos valores de IAT as durações das chamadas deste tipo de evento são na maior parte do tempo as mais longas.

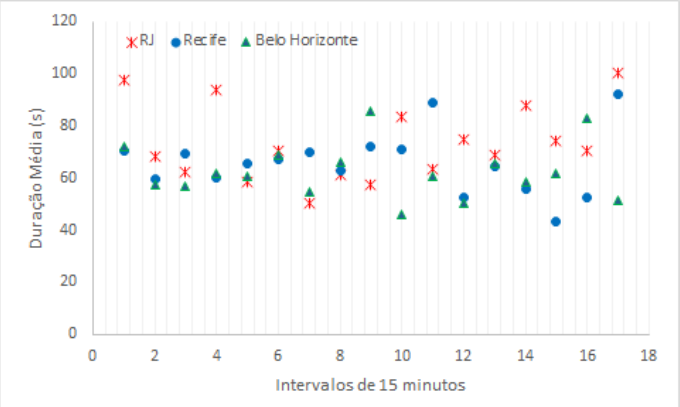


Figura 27 – Duração Réveillon - Belo Horizonte Recife e Rio de Janeiro

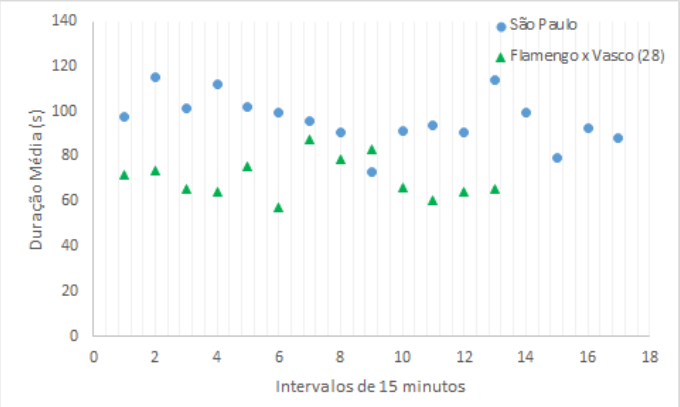


Figura 28 – Comparação Réveillon São Paulo e Flamengo x Vasco (28)

5.3 Análise Comparativa da Região do Evento

A análise das características do evento permite identificar comportamentos específicos de cada tipo de evento, períodos onde os participantes realizam um grande número de chamadas e as características das chamadas, gerando assim informações que possam auxiliar o planejamento da infraestrutura necessária para eventos de larga escala. Em contrapartida, a análise comparativa da região do evento permite descobrir o impacto que um evento de larga escala causa na rede de telefonia móvel que realiza a cobertura da região. As mesmas métricas, utilizadas para analisar as características do evento, serão utilizadas para realizar a comparação entre uma data sem evento na região e uma data com evento.

Tabela 18 – Tabela de duração - Automobilismo

	Duração		
	Duração média (s)	S_{inter}	S_{intra}
GP do Brasil 2012	114,32	36,23	32,35

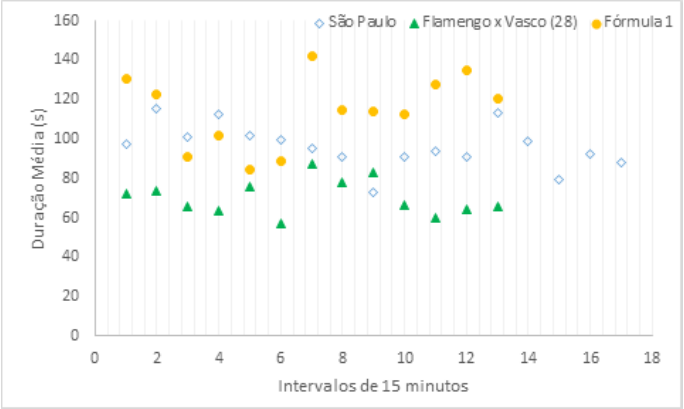


Figura 29 – Duração Réveillon e show

5.3.1 Número de Chamadas

5.3.1.1 Partidas de Futebol

Como esperado, o aumento no número de pessoas na região do evento reflete em volume de chamadas geradas nas ERBs que realizam a cobertura desta área. A Tabela 19 mostra os resultados da análise comparativa para a região do estádio Engenhão. Pode-se notar que nas duas partidas entre Flamengo e Vasco houve um aumento maior que 100% tanto no número de ligações quanto no número de participantes. Entretanto, na partida entre Flamengo e Fluminense os valores foram reduzidos em 75%.

Tabela 19 – Valores totais das bases - Partidas de Futebol

	Chamadas	**Chamadas Sem evento**	**Variação**	**Partic.**	**Partic. Sem evento**	**Variação**
Flamengo x Vasco(28)	11.576	4.256	171%	4.625	2.182	112%
Flamengo x Vasco(04)	14.434	4.256	239%	5.164	2.182	137%
Flamengo x Fluminense	1.055	4.256	-75%	577	2.182	-73%

A Figura 33 mostra a análise agregada, em intervalos de 15 minutos, do número de chamadas na região do estádio Engenhão. Note que na data onde não houve nenhum evento de larga escala o comportamento dos usuários que estiveram na região não segue nenhum padrão. Desta maneira, em alguns períodos, o volume de chamadas desses usuários se aproxima dos valores encontrados quando houve um evento de larga escala, como no jogo entre Flamengo e Vasco, quando o número de usuários na região mais que dobrou.

O planejamento da infraestrutura necessária para eventos de larga escala deve considerar não apenas a natureza do evento, mas também o estudo da região onde o mesmo ocorrerá. Situações como encontradas no estádio Engenhão podem representar cenários onde a instalação de ERBs móveis possam ser necessárias para eventos de grande interesse da população.

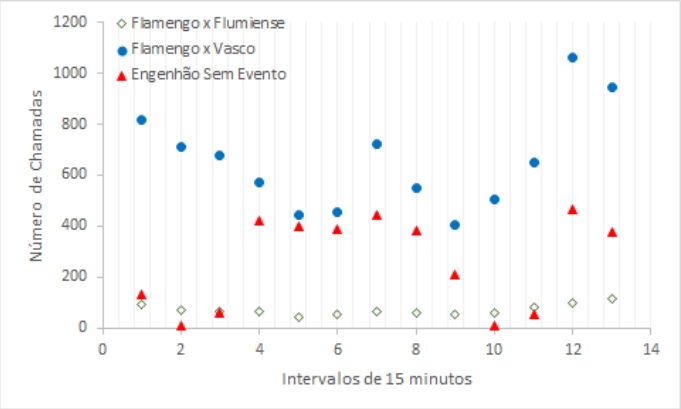


Figura 30 – Análise Comparativa Número de Chamadas - Engenhão

5.3.1.2 Réveillon

A análise das características das festas de Réveillon mostrou a dificuldade de planejar a infraestrutura necessária para suportar esse tipo de evento que recebe uma alta diversidade de pessoas e gera um grande volume de chamadas. Desta maneira, o estudo da região do evento pode auxiliar esse planejamento. A Tabela 20 apresenta valores comparativos entre as festas de Réveillon e a mesma região sem a presença de um evento.

Tabela 20 – Valores totais das bases de dados- Réveillon

	Chamadas	Chamadas Sem evento	Variação	Partic.	Partic. Sem evento	Variação
Belo Horizonte	12.800	5.469	134%	5.121	2.474	106%
Recife	17.016	3.670	364%	5.506	1.508	265%
Rio de Janeiro	14.707	211	6.870%	4.984	74	6.635%
Salvador	28.028	3.176	782%	9.617	1.285	648%
São Paulo	33.474	12.862	160%	11.729	5.815	102%

A coleta da rede de telefonia móvel na região de Copacabana em 03 de janeiro apresentou limitações. A partir do conhecimento da área se pode afirmar que o valor de 74 usuários não representa com fidelidade o comportamento dos usuários que frequentam a região, mesmo sem a existência de um evento de larga escala. Essa afirmação pode ser feita devido ao alto número de hotéis e bares na região, fatores que seriam responsáveis pelo aumento no número de usuários encontrados.

As demais cidades apresentaram aumentos consideráveis tanto em número de chamadas quanto de usuários na região (participantes), sobretudo a cidade de Salvador. Durante a análise de IAT e IDT os resultados indicaram uma possível sobrecarga na rede durante alguns intervalos de 15 minutos. A Figura 31 apresenta o estudo agregado do Réveillon em Salvador e sua respectiva região no dia 03 de janeiro. Percebe-se que nos períodos identificados como possível sobrecarga da rede os valores no número de chamadas são próximos aos identificados no dia 03 de janeiro.

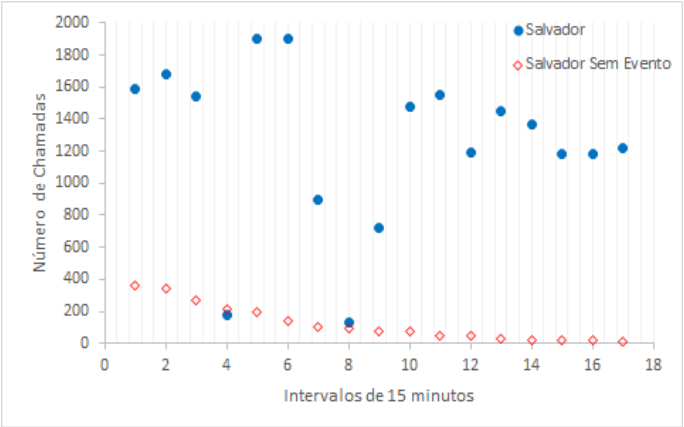


Figura 31 – Análise Comparativa Número de Chamadas - Réveillon de Salvador

5.3.1.3 Automobilismo

Ao contrário das partidas de futebol e das festas de Réveillon, onde em geral o aumento no número de chamadas e participantes foi maior que 100%, durante o Grande Prêmio do Brasil o número de chamadas aumentou 42% e os participantes aumentaram em apenas 25% como mostrado na Tabela 21.

Tabela 21 – Valores totais das bases - Automobilismo

	Chamadas	Chamadas Sem evento	Variação	Partic.	Partic. Sem evento	Variação
GP Brasil 2012	12.308	8.612	42%	4.841	3.871	25%

Uma análise agregada também foi realizada durante o Grande Prêmio do Brasil como mostrado na Figura 32. Durante a análise se percebe que o comportamento da região é constante assim como durante o evento. Entretanto, a análise permitiu identificar um possível pico durante o quarto intervalo. Este intervalo contempla o período onde a largada da corrida aconteceu, possivelmente sendo a responsável pelo pico de chamadas.

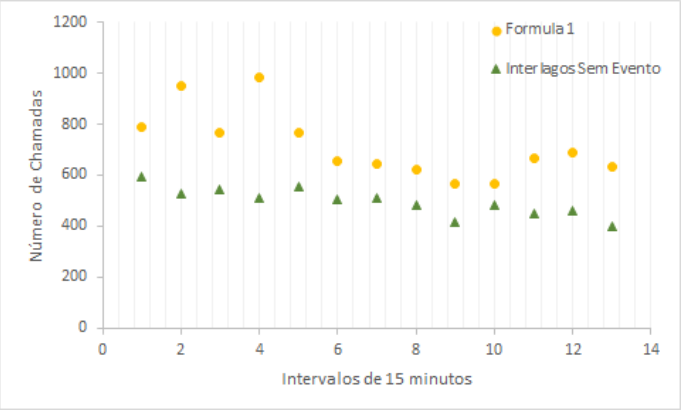


Figura 32 – Análise Comparativa Número de Chamadas - GP do Brasil

5.3.2 Intervalo entre Chamadas

5.3.2.1 Partidas de Futebol

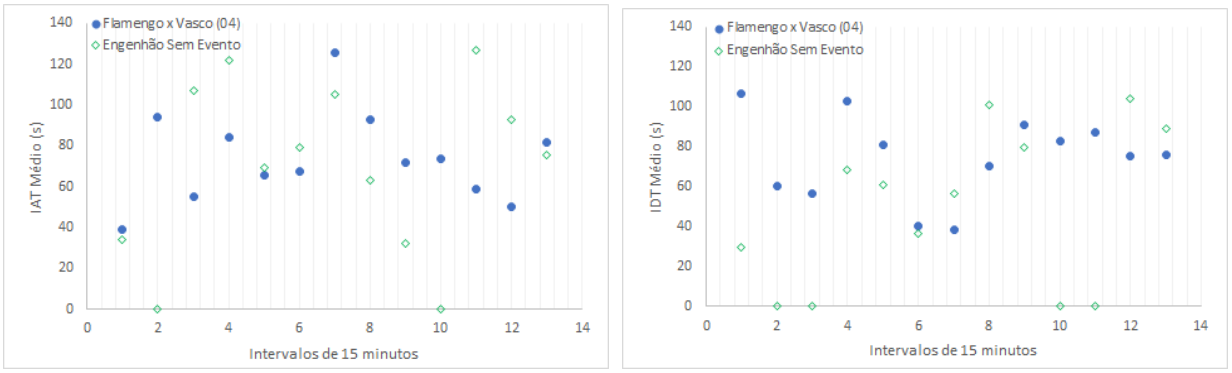
O intervalo médio entre início de chamadas sucessivas e término de chamadas sucessivas durante os jogos do estádio Engenhão apresentaram valores próximos, quando analisada a região sem a ocorrência de um evento. Esses valores podem ser vistos na Tabela 22. A partir desses dados se percebe que o tempo médio de início e término das chamadas sucessivas é próximo mesmo que, em volume, esses valores sejam diferentes. A análise agregada em intervalos de 15 minutos apresentada na Figura 33 confirma que durante todo o evento esses valores são similares.

Tabela 22 – Tabela IAT e IDT - Jogos de futebol

	IAT			**IDT**		
	IAT médio (s)	S_{inter}	S_{intra}	IDT médio (s)	S_{inter}	S_{intra}
Flamengo x Vasco(28)	112,94	116,69	72,44	121,26	128,52	88,63
Flamengo x Vasco(04)	97,12	98,68	57,66	114,44	134,21	53,38
Flamengo x Fluminense	186,95	123,98	124,48	184,47	112,53	103,10
Engenhão sem Evento	107,09	85,44	77,69	115,11	130,74	38,68

5.3.2.2 Réveillon

A análise comparativa do IAT e IDT mostrou que durante o dia 03 de Janeiro essas métricas apresentaram aumentos de até 300% como, por exemplo, em Salvador. Esse comportamento era esperado devido ao grande número de participantes durante as festas de Réveillon nas respectivas regiões. O IAT médio durante as festas apresentou uma redução média de 52%. Este valor mostra que as chamadas chegam com um tempo 52% menor as ERBs. Em Salvador, onde pode ter havido uma sobrecarga da rede, a redução chegou a 73%.



(a) IAT Engenhão sem Evento **(b) IDT Engenhão sem Evento**

Figura 33 – Comparativo Número de chamadas Engenhão Sem evento

A análise mostrou que o tempo que estas chamadas permanecem ocupando recursos da rede também foi reduzido. O IDT médio apresentou redução média de 34%. Como dito anteriormente, valores de IAT maiores que IDT podem levar a uma sobrecarga da rede podendo explicar a dificuldade de realizar ligações durante as festas de Réveillon.

Tabela 23 – Tabela IAT e IDT - Réveillon

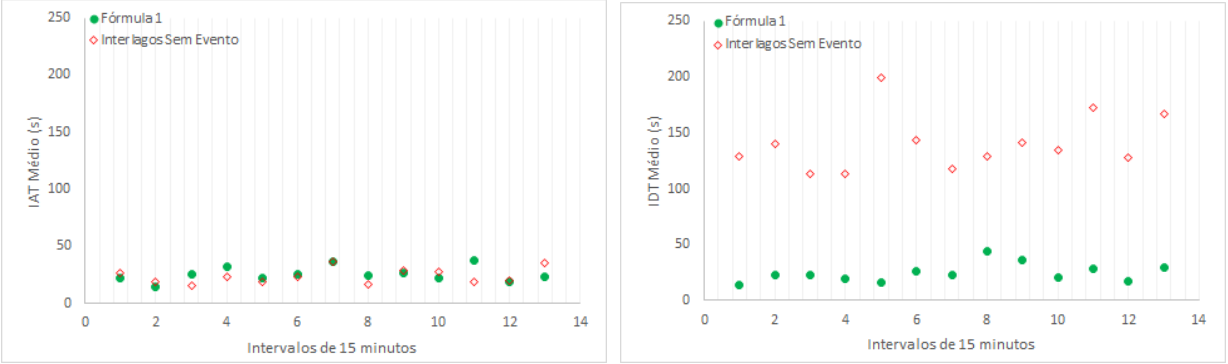
	IAT			IDT		
	IAT médio (s)	S_{inter}	S_{intra}	IDT médio (s)	S_{inter}	S_{intra}
Belo Horizonte	112,38	21,24	25,58	57,03	22,28	58,53
Belo Horizonte Sem Evento	163,99	130,12	95,43	164,94	99,83	110,14
Recife	54,06	74,16	26,23	68,39	72,59	51,51
Recife Sem Evento	108,09	87,41	97,58	117,16	72,09	92,51
Salvador	50,76	70,42	31,00	66,21	104,87	45,04
Salvador Sem Evento	188,71	166,90	140,37	109,51	53,93	71,95
São Paulo	57,39	38,14	29,20	68,66	89,39	30,13
São Paulo Sem Evento	119,30	36,79	95,79	120,41	41,48	105,70

5.3.2.3 Automobilismo

Assim como nas partidas de Futebol, os valores de IAT encontrados durante o Grande Prêmio do Brasil são próximos quando comparados com os dados da região sem a ocorrência de um evento. Entretanto, como mostrado na Tabela 24 os valores médios de IDT são 82% menores. Esse comportamento indica que as chamadas tendem a possuir uma duração menor durante a corrida. O mesmo padrão é identificado durante todo o evento. A Figura 34 permite a visualização agregada, em intervalos de 15 minutos, deste comportamento.

Tabela 24 – Tabela IAT e IDT - Automobilismo

	IAT			IDT		
	IAT médio (s)	S_{inter}	$\overline{S}_{intra}$	IDT médio (s)	S_{inter}	$\overline{S}_{intra}$
GP do Brasil 2012	25,61	40,74	7,03	24,70	38,43	8,77
Interlagos Sem Evento	23,74	33,48	7,44	139,69	41,65	39,29



(a) IAT Interlagos **(b) IDT Interlagos**

Figura 34 – Comparativo IAT e IDT região de Interlagos

5.3.3 Duração

5.3.3.1 Partidas de Futebol

As análises comparativas até então realizadas comprovam que a existência de um evento de larga escala altera a utilização da rede de telefonia móvel da região. A Tabela 25 permite a análise da região do estádio Engenhão sem a existência de um evento. Note que os resultados encontrados mostram um valor de desvio padrão maior que a média.

Tabela 25 – Tabela Comparativa duração - Engenhão Sem Evento

	Duração		
	Duração média (s)	S_{inter}	$\overline{S}_{intra}$
Flamengo x Vasco(28)	70,08	14,54	35,80
Flamengo x Vasco(04)	74,27	15,48	46,59
Flamengo x Fluminense	38,17	12,73	30,53
Engenhão sem Evento	311,84	454,13	446,14

5.3.3.2 Réveillon

Assim como as análises de IAT e IDT, o tempo de duração das chamadas durante as festas de Réveillon apresentou redução. O tempo de duração das chamadas foi reduzido 35% em média. Os valores detalhados de cada festa e sua respectiva região está descrita na tabela 26.

Tabela 26 – Tabela Comparativa duração - Réveillon e Regiões Sem Evento

	Duração		
	Duração média (s)	S_{inter}	$\bar{S}_{intra}$
Belo Horizonte	65,30	22,76	40,35
Belo Horizonte Sem Evento	106,37	38,60	106,85
Recife	58,24	18,94	22,65
Recife Sem Evento	86,49	14,78	47,54
Salvador	67,39	27,91	35,89
Salvador Sem Evento	122,69	61,82	145,51
São Paulo	98,03	29,41	40,84
São Paulo Sem Evento	126,00	32,36	102,78

A Figura 35 mostra um comparativo entre o Réveillon de Belo Horizonte e outras três regiões onde ocorreram comemorações. Percebe-se que apenas no final do evento a duração das chamadas apresenta valores próximos ao encontrados na região sem a presença de um evento. Durante a maior parte do período analisado, as chamadas realizadas durante as festas de Réveillon tendem a ter uma duração menor se comparadas com suas respectivas regiões sem evento.

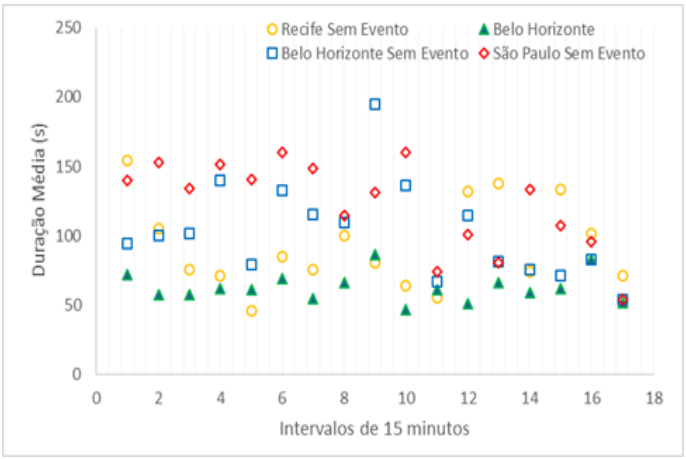


Figura 35 – Duração Réveillon e show

5.3.3.3 Automobilismo

Assim como nos demais eventos, a análise da corrida automobilística mostrou que durante a realização do evento os participantes tendem a realizar chamadas com menor duração. A redução encontrada foi de 19% e os valores são descritos na tabela 27

Tabela 27 – Tabela Comparativa duração - Região de Interlagos

	Duração		
	Duração média (s)	S_{inter}	S_{intra}
GP do Brasil 2012	114,32	36,23	32,35
Interlagos Sem Evento	140,62	40,85	40,17

5.4 Análise da Mobilidade

Para analisar a dinâmica da carga nas ERBs localizadas nas rotas de acesso ao local do evento são utilizados mapas de calor para representar a intensidade de utilização de cada ERB pelo serviço de telefonia móvel. Os mapas são gerados por uma função da Google Maps Javascript API V3 (MAPS, 2012), que recebe como entrada a latitude, a longitude e o número de ligações de cada ERB analisada.

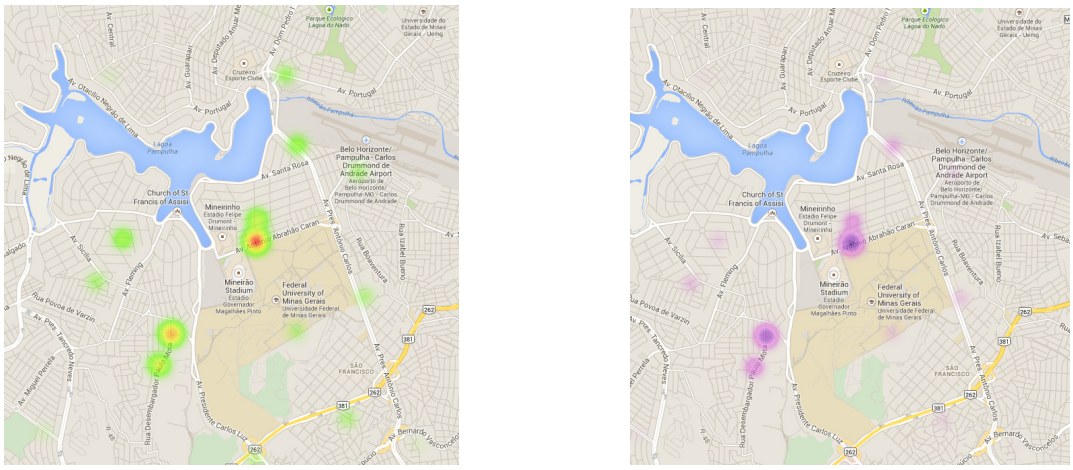
A metodologia de mobilidade foi criada com o intuíto de possibilitar uma análise mais aprofundada da movimentação dos usuários da rede de telefonia móvel ao decorrer do tempo. A principal utilização dos dados após a execução da metodologia é na geração dos vídeos da mobilidade dos usuários. Entretanto, a análise dos períodos $t_{chegada}$ e $t_{partida}$ pode prover informações importantes tanto para o planejamento das cidades como das operadoras de telefonia.

As análises geradas nas últimas seções mostrou com detalhes o comportamento dos participantes na região do evento, porém, é importante conhecer quais as regiões que esses participantes se originam, para onde vão, quais vias utilizam entre outras informações. Nesta seção serão realizadas as análises que permitam apresentar respostas para essas e outras perguntas. Como dito anteriormente todos os mapas gerados para a análise utilizam da Google Maps Javascript API V3 (MAPS, 2012). O processo de criação dos mapas será mais bem detalhado na seção 6.

Diferente das análises de caracterização da região do evento a análise comparativa da mobilidade dos participantes depende não apenas do tipo do evento, mas também do local. A forma que os usuários acessam o local do evento varia de acordo com a cidade e localização. Desta maneira a análise comparativa será realiza com eventos que ocorreram no mesmo local independente do seu tipo (futebol, show ou Réveillon). Note que, a análise da mobilidade continua levando em consideração apenas os espectadores do evento.

O estádio do Mineirão em Belo Horizonte recebeu partidas de futebol e o show do Paul McCartney. Sendo o único local a receber eventos de natureza distinta e possuindo coletas sem perdas de dados, as análises comparativas da mobilidade dos usuários serão realizadas utilizando essas bases de dados. Outro fator determinante para a escolha destas bases de dados está no amplo conhecimento do autor sobre a região do estádio.

As partidas entre Cruzeiro e Atlético e entre Brasil e Chile, após a análise da caracterização dos dados, podem ser consideradas como partidas festivas. Entretanto, pequenas diferenças podem ser notadas quando se analisa a mobilidade dos seus participantes. A Figura 36 mostra como os participantes se movimentaram durante o período do t_{evento}. Pode-se notar que na partida entre Brasil e Chile a concentração das chamadas acontece principalmente na ERB que cobre o estádio. Entretanto, na partida entre Cruzeiro e Atlético pontos na região próxima ao estádio podem ser notados. Nota-se que alguns destes pontos também aparecem na Figura 36b porém com uma intensidade menor.



(a) Mobilidade Cruzeiro x Atlético **(b) Mobilidade Brasil x Chile**

Figura 36 – Mobilidade participantes durante o t_{evento} no Mineirão

Sabe-se que o comportamento dos participantes da partida entre Cruzeiro e Atlético é bastante característico. Nesta partida de futebol os torcedores tendem a chegar mais cedo ao estádio e permanecem nos bares próximos ao estádio até a proximidade do horário da partida. Para garantir um serviço de qualidade aos clientes as operadoras devem concentrar seu planejamento não apenas no local onde o evento será realizado, mas também em todo seu entorno. Esta análise permite conhecer o perfil dos usuários nestes eventos de natureza similar, porém de características distintas quando analisada a mobilidade dos usuários.

A utilização das vias é outro conhecimento passível de ser retirado das análises da mobilidade dos participantes em eventos de larga escala. Dentre todas as possíveis rotas de acesso e saída do estádio Mineirão, pode-se determinar a Avenida Presidente Antônio Carlos como a principal. A Figura 37 mostra o mapa da mobilidade dos usuários das partidas entre Cruzeiro e Atlético entre Brasil e Chile, além do show do músico Paul McCartney. Como esperado os participantes tendem a utilizar a Avenida Antônio Carlos, pois durante toda a via as ERBs que a cobrem recebem um valor considerável de chamadas.

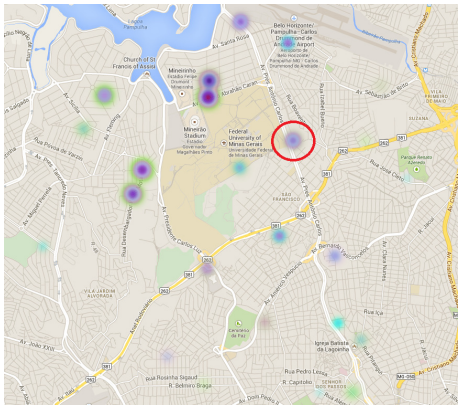


Figura 37 – Mobilidade $t_{partida}$ Cruzeiro x Atlético (Verde) - Brasil x Chile (Roxo) - Show Paul McCartney (Azul)

Uma análise mais detalhada da Figura 37 permite identificar uma região da avenida, onde durante a saída de grandes eventos constantemente gera-se um congestionamento de veículos. Na região destacada em vermelho no mapa, mesmo consideravelmente distante do estádio, devido à existência de um semáforo é criado um pequeno congestionamento de veículos. Em todos os três eventos analisados essa região apresentou um grande número de chamadas durante o $t_{partida}$ o que permite inferir que enquanto estão parados no trânsito os participantes tendem a utilizar dos recursos da rede.

6 PREDITRAF

O *PrediTraf* é um sistema de identificação da localização do uso concentrado da rede móvel em áreas e vias públicas e privadas. Utilizando de dados retirados das CDRs e baseado nas metodologias descritas no Capítulo 4, o *PrediTraf* foi desenvolvido para possibilitar uma análise do uso da rede de telefonia móvel em diversos cenários. Acredita-se que, em trabalhos futuros, o *PrediTraf* possa evoluir ao ponto de se tornar um simulador. Desta maneira a ferramenta foi projetada seguindo os modelos descritos em Jain (1991) para desenvolvimento de simuladores.

Como descrito em Jain (1991), grande parte das ferramentas desenvolvidas produzem resultados insatisfatórios devido à baixa preocupação com o suporte estatístico ou com seu desenvolvimento. Para evitar tais problemas o desenvolvimento desses módulos fora realizado separadamente. Além dos módulos estatístico e de execução, o *PrediTraf* possui pacotes específicos para geração de mapas, vídeos e gráficos que serão detalhados nas seções a seguir.

Segundo a literatura utilizada um passo importante do desenvolvimento de uma ferramenta é a definição da linguagem adotada. Essa definição consiste na escolha de uma linguagem de simulação ou uma linguagem de propósito geral. As linguagens de simulação podem ser divididas em duas categorias: linguagens de simulação contínuas ou linguagens de simulação por eventos discretos. Todavia, ambas as categorias possuem uma característica em comum: a alta inflexibilidade. Devido a essa característica e juntamente com a necessidade de análises de eventos distintos, foi definido o uso de uma linguagem de propósito geral. Sendo assim, a linguagem definida para o desenvolvimento foi o JavaFX.

A linguagem JavaFX é uma plataforma de software disponibilizada pela Oracle que permite a criação e disponibilização de aplicações para diversos ambientes. O desenvolvimento segue o padrão MVC (Model, View and Control) e permite a utilização de classes e bibliotecas JavaSE. A possibilidade de integração com os mapas gerados pela Google Maps Javascript API V3 Maps (2012) de forma inerente ao sistema foi um dos fatores determinantes para a escolha, além da portabilidade comum de todas as aplicações Java.

6.1 O Modelo do PrediTraf

6.1.1 Modelo de Simulação do PrediTraf

Em Jain (1991) são descritas 14 características de um modelo de simulação. Dessas 14, o *Preditraf* apresenta cinco características, listadas na Tabela 28.

Tabela 28 – Tabela de Características de Modelos de Simulação

Modelo	Características
Estado Discreto	Modelo onde as variáveis são definidas em tempos pré-determinados
Determinístico	Modelo onde os resultados podem ser previstos caso o mesmo conjunto de entradas seja utilizada.
Estático	Modelo onde o tempo não é considerado uma variável.
Linear	Modelo onde o número de entradas segue uma função linear com o número de saídas
Aberto	Modelo onde as entradas podem vir de locais externo ao sistema

O *PrediTraf* foi desenvolvido utilizando características de diferentes modelos. Possuindo um módulo de probabilidade, mesmo que básico, podemos definir que o *PrediTraf* possui características determinísticas. Essas características existem, pois, as funções de probabilidade são utilizadas apenas para simular o aumento da carga, fazendo com que as análises possam ser repetidas. As características de um modelo estático também são encontradas na ferramenta. Apesar do tempo ser uma variável de entrada, ele é apenas considerado para delimitar a janela de coleta dentro da base de dados de referência. Podemos definir também a ferramenta como linear, pois, as seis variáveis de entrada são responsáveis pela geração das cinco saídas, e aberto, por utilizar de dados gerados externamente à ferramenta. As entradas juntamente com a base de dados de referência serão detalhadas nas seções a seguir.

6.1.2 Entradas do Modelo

Após a definição das características do modelo a ser desenvolvido, o próximo passo tomado foi a seleção das variáveis de entrada e saída do mesmo. Inicialmente foram levantadas 26 variáveis de entradas para o modelo. Entretanto, segundo Jain (1991), um dos principais problemas durante o desenvolvimento de uma ferramenta está na alta complexidade dos modelos definidos. Um grande número de variáveis de entrada aumenta o tempo de desenvolvimento e a probabilidade de erros durante a análise. Para solucionar esse problema foram listadas as variáveis essenciais para a realização das análises e posteriormente acrescentadas àquelas necessárias para obtenção dos resultados desejados. Após o processo de seleção definiu-se como seis as variáveis principais do sistema:

1. Horário de início do evento
2. Horário de do evento
3. Público total esperado
4. *Market Share* do local
5. Horário de chegada dos usuários
6. Horário de partida dos usuários

As duas primeiras variáveis (horário de início do evento e horário de término do evento) são responsáveis pela definição da linha de tempo descrita no Capítulo 4. Os horários definidos por essas duas variáveis serão utilizados para estabelecer o t_{evento}. A partir da definição do período do t_{evento} será somado ou subtraído os respectivos tempos do $t_{chegada}$ e $t_{partida}$, definindo assim o tamanho da linha de tempo. As outras quatro variáveis são utilizadas para a execução da metodologia de mobilidade descrita na Seção 4.2.

O público esperado e o *market share* definem o número de usuários que as ERBs que cobrem a região do evento terão que atender. Tanto o público esperado e o *market share* são informações previamente conhecidas, sendo que a segunda pode ser acessada no site da Agência Nacional de Telecomunicações (ANATEL*). As duas últimas variáveis, horário de chegada dos usuários e horário de partida dos usuários, são utilizadas para aplicação da metodologia de mobilidade para os novos usuários calculados a partir do público e do *market share*.

Entre as variáveis anteriormente levantadas e não utilizadas na definição do modelo do *PrediTraf*, algumas dessas não foram utilizadas como entrada por fazerem parte do processo de análise. As demais não estão presentes na primeira versão da ferramenta, pois, gerariam um grande aumento de complexidade e serão apresentadas em trabalhos futuros.

A ferramenta deve possibilitar que análises anteriores sejam carregadas. Para isso, existe a possibilidade no modelo da entrada de dados via arquivos, previamente gerados e salvos pela ferramenta. Os arquivos gerados possuem os dados da execução das metodologias de caracterização e mobilidade.

Segundo Jain (1991) um *race* é uma lista de eventos ordenados pelo tempo e suas variáveis associadas, podendo apresentar vários níveis de detalhamento. Para o uso da ferramenta *PrediTraf* é necessária a existência de uma base de dados de referência. Os dados presentes nessa base de dados foram gerados a partir de um arquivo contendo o

*www.anatel.gov.br

detalhamento de cada chamada realizada durante uma determinada janela de tempo. Os dados necessários para a inserção na base de dados são:

- Identificador único da chamada
- Código do país
- DDD
- Setor de Início da chamada
- Setor de Término da chamada
- Número de *Handovers*
- Data de início da chamada
- Horário de inicio da chamada
- Data de fim da chamada
- Horário de fim da chamda
- Duração da chamada

6.1.3 Saídas do Modelo

A partir das variáveis de entrada e aplicando as metodologias de caracterização e mobilidade, o modelo permite a geração de cinco saídas. A Figura 38 mostra o modelo do *PrediTraf* com suas entradas e saídas.

Os resultados gerados pelo modelo podem ser divididos em duas partes. O mapa geral de carga, o relatório de impactos e os gráficos de ocupação são responsáveis por mostrar os resultados da metodologia de caracterização. As demais saídas são responsáveis por exibir os resultados da metodologia de mobilidade.

6.2 Arquitetura do PrediTraf

Após a definição do modelo a ser desenvolvido com suas características e variáveis de entrada e saída, o próximo passo do desenvolvimento da ferramenta foi a definição da arquitetura para desenvolvimento. A Figura 39 mostra o diagrama de pacotes da arquitetura.

A arquitetura do *PrediTraf* pode ser dívida em quatro áreas diferentes. Os pacotes dentro das áreas de cor azul (*Views* e *Controllers*) possuem as classes responsáveis pela

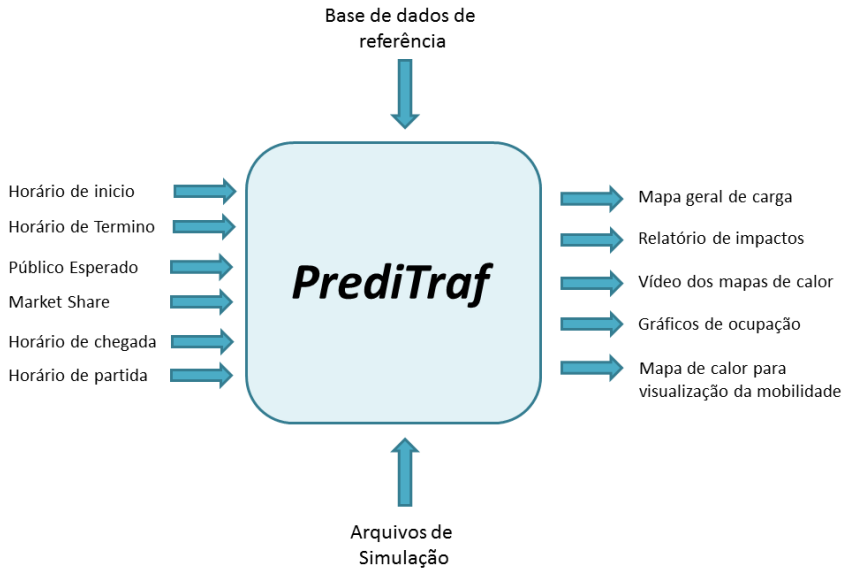


Figura 38 – Modelo da ferramenta

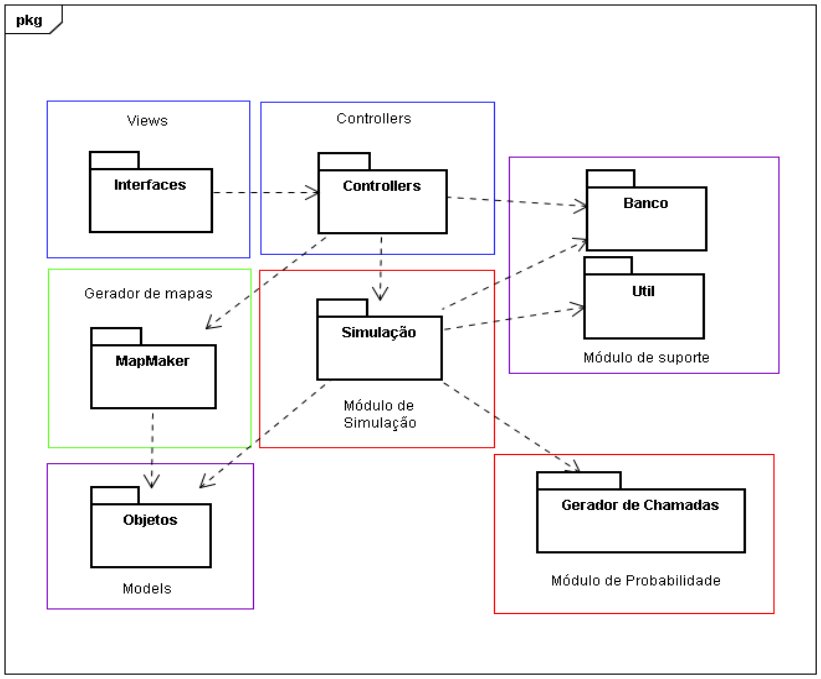


Figura 39 – Diagrama de pacotes da ferramenta

gestão da interface e comunicação com o módulo de simulação. Essas classes seguem o padrão de desenvolvimento da linguagem JavaFX. Os pacotes dentro das áreas em vermelho compõem os módulos que executam a análise do evento (Módulo de simulação e Módulo de probabilidade). Dentro destes módulos foi realizada a implementação das metodologias.

Os pacotes pertencentes às áreas de cor roxa possuem as classes de suporte a ferramenta. Dentro do pacote de objetos estão as definições de ERBs, local do evento, chamada e seus respectivos componentes. Os pacotes do módulo de suporte possuem as classes responsáveis pela comunicação com o banco de dados e escrita dos arquivos das análises. A área em verde contém o módulo de geração de mapas. Neste pacote estão implementadas as classes que utilizam da Google Maps Javascript API V3 Maps (2012) para geração dos mapas utilizados, além das classes responsáveis pela geração do vídeo. Devido à alta complexidade esse pacote será melhor detalhado na próxima seção.

6.2.1 Módulo de Geração de Mapas e Vídeos

Para visualizar o dinamismo da carga de trabalho de um dado evento e verificar a concentração de pessoas de forma temporal, foi desenvolvido uma forma de aperfeiçoar a geração dos mapas de calor. Tal automatização permite conseguir a representação da intensidade na utilização da rede de telefonia móvel ao longo do tempo em que o evento acontece.

Para a geração dos mapas de calor, foi utilizada a API do Google Maps Javascript V3, mais especificamente a visualização em forma de *Heat Maps* (Mapas de Calor). Essa API oferece diversos utilitários para manipulação e adição de conteúdo aos mapas. Sendo um serviço gratuito oferecido pela Google, a forma de visualização utilizada permite a geração de mapas de calor baseados em uma escala de cor nas respectivas localizações das ERBs.

Trabalhando com eventos de duração cada vez maiores, como um Réveillon, a automação na geração dos mapas de calor se tornou essencial para o desenvolvimento do trabalho. Tais eventos necessitavam da geração de até 300 mapas, tornando assim o modelo de criação manual inviável. Com esta melhoria, se pode obter vários mapas simultaneamente, trazendo assim resultados mais detalhados não só de forma visual, mas temporal.

Para criação automática dos mapas foi desenvolvida, em linguagem Java, uma rotina que recebe os seguintes parâmetros:

- As coordenadas geográficas (latitude e longitude) das ERBs da região do evento;
- Vetor contendo o número ligações realizadas nestas ERBs;
- O horário de início da análise;

Após a entrada dos parâmetros ocorre a geração de arquivos HTML para cada minuto do evento, onde a própria API do Google é responsável pela distribuição de cores.

Esta distribuição obedece regras definidas de acordo com a concentração no número de chamadas, em determinada região e ERB, além do tempo com que essa concentração acontece.

Após a criação dos arquivos HTML, contendo os mapas de calor característicos minuto a minuto de dado evento, esses arquivos são enviados para o freeware SiteShoter para obtenção da imagem dos mapas (NIR SOFER, 2011). O SiteShoter é um software que possibilita a gravação de imagens de quaisquer páginas web em arquivo, esse utilitário permite também ao usuário converter arquivos HTML em imagem. Neste módulo do *PrediTraf*, o SiteShoter é utilizado para converter os mapas de calor (arquivo HTML) em frames de vídeos de calor (arquivo de imagem JPEG).

O SiteShoter possui uma interface do usuário. Porém com intuito de tornar o processo de geração de vídeos automático é utilizado o modo de linha de comando, onde o mesmo é invocado pela própria rotina em Java. Após sua chamada, automaticamente o sistema converte os arquivos HTML criados em arquivos de imagem JPEG.

A geração de imagens é o processo mais demorado na geração dos vídeos de calor, uma vez que cada frame do vídeo leva cerca de 8 segundos para ser obtido. Como o processo de obtenção de imagens se resume na captura da tela do arquivo HTML, sem a necessidade abrir um browser, caso ocorra uma falha de conexão durante este processo de captura, a qualidade do vídeo ficará comprometida uma vez que haverá frames sem a imagem correta.

Para melhor visualizar a carga das ERBs, no período de tempo em que o vídeo é exibido, foi desenvolvida uma função em Java responsável por inserir um pequeno relógio no canto inferior direito da tela e uma escala de cor contendo o valor máximo de chamadas no canto superior direito da tela. A Figura 40 apresenta um frame tendo o relógio e a escala de cor inseridos.

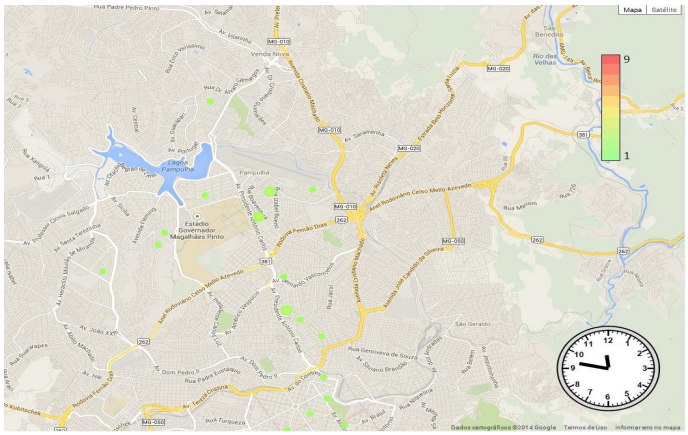


Figura 40 – Frame do vídeo de mobilidade

O relógio é inserido de acordo com o horário real em que ocorreu a utilização da ERB e com o decorrer do tempo seu valor é alterado. Já a escala é calculada a partir do valor máximo de ligações de uma ERB. O valor máximo encontrado em uma ERB coletada passa a ser o valor máximo da escala. Os demais valores vão sendo calculados por correspondência. Esse processo de inserção de duas imagens (relógio e escala) nos *frames* levam alguns milissegundos por imagem, não comprometendo o tempo gasto para geração dos vídeos de calor.

A montagem do vídeo em Java é efetuada através da utilização da biblioteca Xuggler (CONNECTSOLUTIONS, 2011). O Xuggler é uma biblioteca de código aberto para desenvolvedores Java utilizada para descompactar, manipular e comprimir um vídeo gravado ou ao vivo.

A função em Java utilizando Xuggler é capaz de, a partir de um diretório contendo arquivos de imagens, concatena-las criando um vídeo e salvando-o na máquina do usuário. São utilizados valores padrões para as taxas de FPS (Frames por Segundo) e de codificação do vídeo. O formato de saída do vídeo é um arquivo com extensão MP4 contendo a visualização espaço-temporal do evento.

É necessário que todos os frames do vídeo estejam prontos antes de serem utilizados na montagem do vídeo. Sendo assim, a função de criação do vídeo só é iniciada após a conversão de todas as imagens HTML em JPEG, e inserção do relógio e da escala nas mesmas.

A geração dos arquivos HTML, a conversão em imagens, a inserção do relógio escala e por último a geração do vídeo a partir dos frames é imperceptível pelo usuário. A função Java é responsável por salvar e manipular os arquivos em determinadas pastas do sistema. Essas pastas são criadas e apagadas ao fim de cada execução, sendo resultante somente o vídeo de calor ao final.

6.3 Interfaces

6.3.1 Interface Principal

A aplicação desenvolvida a partir da arquitetura, anteriormente detalhada, pode ser dividida em três partes. O módulo de carga, a interface de resultados e a interface principal que será descrita nesta seção. A Figura 41 mostra a interface inicial do *PrediTraf*.

O primeiro passo para utilização da ferramenta é definir qual a base de dados de referência será utilizada durante a análise. Para tal definição é necessário informar a cidade a qual essa base pertence, o tipo de evento que foi coletado e o local do mesmo. Após a escolha da base de referência uma análise prévia dos dados é realizada. Nessa análise

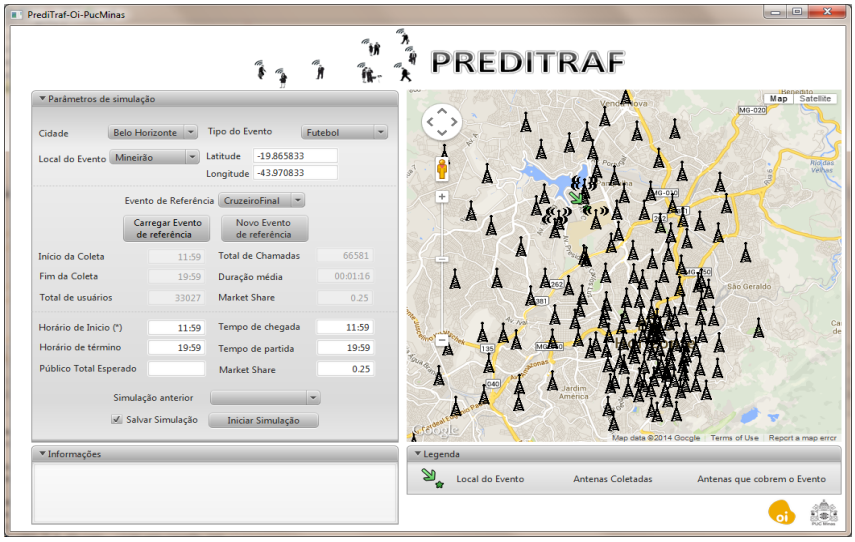


Figura 41 – Tela inicial do PrediTraf

são levantados alguns dados que auxiliem o usuário a definir os parâmetros de entrada da ferramenta. Os dados pré-carregados informam o total de chamadas e usuários da base de dados, o horário de início e fim da coleta e a duração média das chamadas. Além destes dados de auxílio, a ferramenta executa uma avaliação inicial das ERBs. Essa avaliação permite visualizar a atual situação de ocupação das ERBs, ou seja, se existem setores já sobrecarregados. Essa visão geral da rede de telefonia móvel permite uma comparação com os dados gerados após a análise.

Com a avaliação inicial da base de dados de referência, o passo seguinte consiste na inserção das variáveis de entrada da ferramenta. Como descrito no modelo, as seis variáveis de entrada são essenciais para a execução da análise do evento de larga escala. Existe ainda a possibilidade de carregar análises anteriores. Essa funcionalidade se mostrou necessária para otimizar o tempo gasto para análise de eventos de longa duração. Ao selecionar uma análise anterior os arquivos gerados durante a mesma são automaticamente carregados. As variáveis definidas na análise pré-carregada também são preenchidas com a exceção do público esperado. Essa decisão foi tomada por acreditar que essa seja a variável com maior possibilidade de variação.

6.3.2 Interface de Resultados

Após a finalização dos cálculos os resultados obtidos são exibidos na interface de resultados. A Figura 42 exibe um exemplo da interface após a realização de uma análise.

Para utilização da interface de resultados, inicialmente se verifica o mapa gerado para visualizar a situação das ERBs após a análise. O mapa gerado permite identificar regiões que houve uma grande utilização na rede de telefonia móvel. Identificadas as

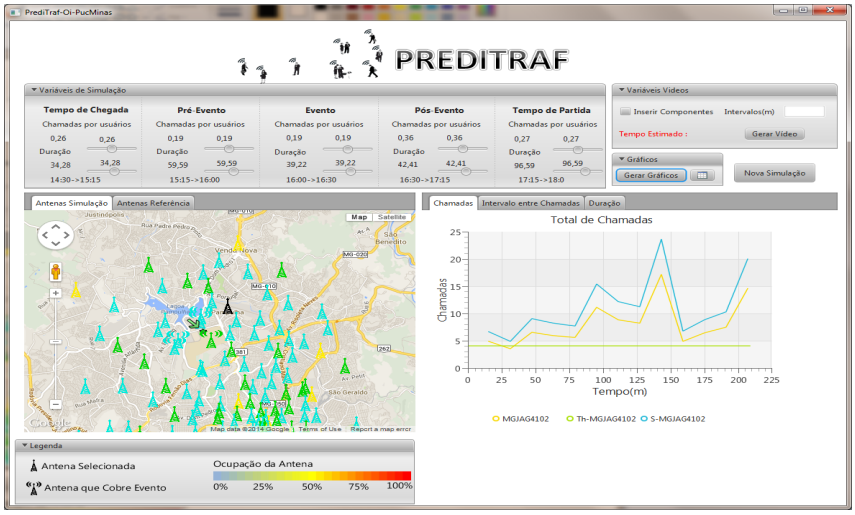


Figura 42 – Tela de resultados do PrediTraf

ERBs ou regiões que se deseja uma informação detalhada sobre sua utilização, se pode selecionar as ERBs para geração dos gráficos de utilização. Após a geração dos gráficos a aba contendo as variáveis de análise tem os valores médios das chamadas por usuário e duração preenchidos. Os valores são preenchidos sobre a divisão de tempo estabelecido na linha de tempo da metodologia de caracterização, ou seja, $t_{chegada}$, *Pré-Evento*, *Evento*, *Pós-Evento* e $t_{partida}$. A alteração desses valores sobre os intervalos da linha de tempo permite uma customização total do evento.

A segunda análise que a interface de resultados permite realizar é sob o aspecto da mobilidade. Os resultados da metodologia de mobilidade são apresentados em forma de mapas de calor. Os mapas permitem ao usuário da ferramenta uma visualização de como os participantes estão se locomovendo antes, durante e após o evento. Essa informação permite que se identifiquem locais onde os participantes ficam presos em engarrafamentos ou para onde se locomovem após o término do evento. Tal dado pode auxiliar a operadora a definir estratégias para prover um serviço de qualidade aos usuários também nesses locais. O local e horário do evento é de conhecimento comum. Sendo assim, se preparar para suportar o número de usuários novos que irão a esse local não exige um estudo da mobilidade. Entretanto, os locais onde os usuários se locomovem após o evento ou onde eles ficam concentrados devido a um engarrafamento, podem variar de acordo com o evento. Figura 43 apresenta um exemplo de mapas de calor gerado pelo *PrediTraf*.

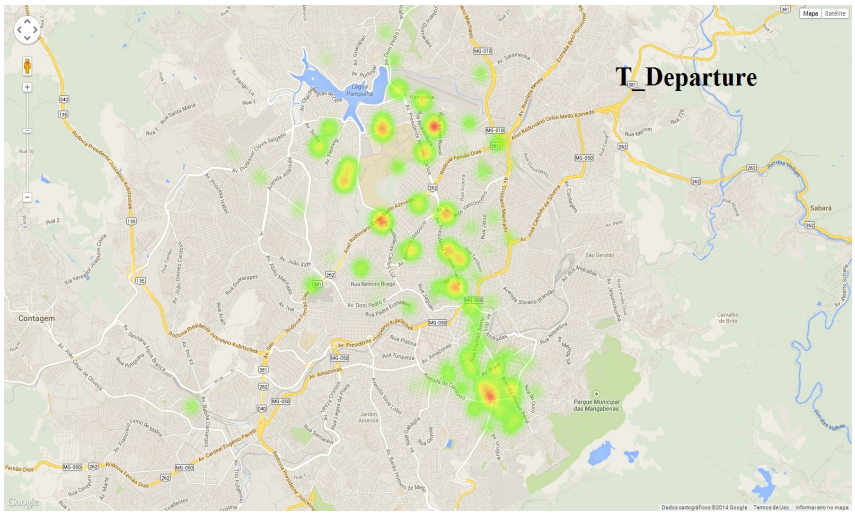


Figura 43 – Mapa de calor da mobilidade dos usuários

7 CONCLUSÃO

A caracterização eficiente dos participantes de grandes eventos pode auxiliar o planejamento da infraestrutura necessária para atendê-los. Este trabalho apresentou características específicas desses usuários em eventos de natureza distintas. Análises inéditas de novas bases de dados, como os jogos da Copa do Mundo de 2014 e do Grande Prêmio do Brasil, permitiram entender melhor o comportamento dos participantes destes e dos demais eventos. Essas análises de diversas bases de dados contendo um grande volume de informação pode caracterizar este trabalho como uma análise de *Big Data*. Este trabalho levou em conta problemas comuns na análise de *Big Data* como heterogeneidade, escalabilidade, oportunidade, complexidade e de privacidade, listados em Labrinidis e Jagadish (2012), apresentando resultados que permitem o avanço computacional e o aumento da rentabilidade de muitas empresas

O estudo de diversas bases de dados possibilitou a definição do efeito "W", comportamento característico entre os participantes de jogos de futebol. Entre os principais resultados apresentados está a comprovação que não apenas o número de chamadas mas também a duração segue uma lei de potência em determinados eventos de larga escala. Esse estudo da duração estende os resultados encontrados em Tanahashi et al. (2012) e contribui aos resultados encontrados em Song et al. (2010b), quando apresenta mais uma característica que pode aumentar a dificuldade de se analisar a mobilidade dos usuários a partir de dados da telefonia celular.

Sendo um problema global, o planejamento da rede para eventos de larga escala pode ser auxiliado por meio do uso de ferramentas como o *PrediTraf*. O Brasil sediou a Copa do Mundo de 2014 e sediará as Olimpíadas de 2016. Esses são os dois maiores eventos esportivos do mundo. Desta forma, analisar o comportamento dos usuários nestes eventos pode auxiliar os próximos países a se prepararem de forma adequada. A análise dos participantes dos jogos da Copa do Mundo mostrou que o efeito "W" é encontrado também nessas partidas de futebol e deve ser considerado durante o planejamento da infraestrutura da rede de telefonia móvel do evento.

Os resultados da análise comparativa da região do evento confirmaram que, apesar da dificuldade, o planejamento da infraestrutura da rede de telefonia móvel para grandes eventos deve ser realizado de forma adequada. O aumento no número de usuários da região em até 160%, como por exemplo no Réveillon de São Paulo, mostra que eventos de larga escala exercem alterações no uso da rede de telefonia móvel na região do evento.

Este trabalho mostrou que a análise agregada em intervalos de tempo de diversas métricas possibilita encontrar diversas situações ao longo do evento. As métricas de IAT e IDT além do número de chamadas e duração podem auxiliar as empresas de telefonia a identificarem regiões críticas durante eventos de larga escala. Como mostrado durante a análise do Réveillon de Salvador, a relação entre o número de chamadas e o tempo entre início de chamadas sucessivas permite identificar períodos que indiquem uma alteração na rede de telefonia móvel.

Este trabalho apresenta uma pesquisa com grandes possibilidades de evolução, como por exemplo, a análise de outras bases de dados referentes a shows e o estudo se as métricas de IAT e IDT também seguem uma lei de potência. Caracterizar de forma eficiente os participantes desse tipo de evento ainda é necessário para confirmação da sua forma de uso da rede de telefonia móvel. Aplicar as metodologias de caracterização e mobilidade em novas bases de dados contendo informações acerca do uso de SMS e 3G/4G é essencial para uma visão completa do uso dos recursos disponíveis na rede de telefonia móvel. Os resultados gerados pela aplicação das metodologias nessas bases de dados, juntamente com as informações sobre o uso da telefonia móvel, irá possibilitar encontrar um número maior de participantes e consequentemente melhores resultados. Desta maneira, se faz necessária que a ferramenta *PrediTraf* evolua para suportar a inserção de dados coletados tanto de SMS quanto da rede 3G/4G. O desenvolvimento de um módulo estatístico avançado que permita o *PrediTraf* realizar simulações das regiões analisadas também será contemplado nos trabalhos futuros.

REFERÊNCIAS

ASGARI, F.; GAUTHIER, V.; BECKER, M. A survey on human mobility and its applications. CoRR, abs/1307.0814, 2013.

BAGROW, J. P.; WANG, D.; BARABáSI, A.-L. Collective response of human populations to large-scale emergencies. PLoS ONE, Public Library of Science, v. 6, n. 3, p. e17680, 03 2011.

BALCAN, D. et al. Multiscale mobility networks and the spatial spreading of infectious diseases. Proceedings of the National Academy of Sciences, 2009. Disponível em: <http://www.pnas.org/content/early/2009/12/11/0906910106.abstract>.

BATTY, M. et al. Smart cities of the future. The European Physical Journal Special Topics, Springer-Verlag, v. 214, n. 1, p. 481–518, 2012. ISSN 1951-6355. Disponível em: <http://dx.doi.org/10.1140/epjst/e2012-01703-3>.

BATTY, M.; DESYLLAS, J.; DUXBURY, E. The discrete dynamics of small-scale spatial events: agent-based models of mobility in carnivals and street parades. International Journal of Geographical Information Science, v. 17, n. 7, p. 673–697, 2003. Disponível em: <http://dx.doi.org/10.1080/1365881031000135474>.

BLEICHER, A. The on-demand olympics. IEEE Spectrum, v. 49, n. 7, p. 9–10, jul 2012.

BROCKMANN, D.; HUFNAGEL, L.; GEISEL, T. The scaling laws of human travel. Nature, v. 439, n. 7075, p. 462–465, jan. 2006. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/16437114>.

CACERES, N.; WIDEBERG, J. P.; BENITEZ, F. G. Deriving origin destination data from a mobile phone network. Intelligent Transport Systems, IET, v. 1, n. 1, p. 15–26, March 2007. ISSN 1751-956X.

CALABRESE, F. et al. The geography of taste: Analyzing cell-phone mobility and social events. In: Proceedings of the 8th International Conference on Pervasive Computing. Berlin, Heidelberg: Springer-Verlag, 2010. (Pervasive'10), p. 22–37. ISBN 3-642-12653-7, 978-3-642-12653-6.

CANDIA, J. et al. Uncovering individual and collective human dynamics from mobile phone records. Journal of Physics A: Mathematical and Theoretical, IOP Publishing, v. 41, n. 22, p. 224015, 2008.

CHO, E.; MYERS, S. A.; LESKOVEC, J. Friendship and mobility: User movement in location-based social networks. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA: ACM, 2011. (KDD '11), p. 1082–1090. ISBN 978-1-4503-0813-7. Disponível em: <http://doi.acm.org/10.1145/2020408.2020579>.

CLAUSET, A.; SHALIZI, C. R.; NEWMAN, M. E. J. SIAM Review, v. 51, n. 4, p. 661–703, 2009. Disponível em: <http://dx.doi.org/10.1137/070710111>.

CONNECTSOLUTIONS. Xuggler. Disponível em <http://www.xuggle.com/>. Acesso em: 05 de out. 2014. 2011.

EAGLE, N.; PENTLAND, A. S.; LAZER, D. Inferring friendship network structure by using mobile phone data. Proceedings of the National Academy of Sciences, 2009. Disponível em: <http://www.pnas.org/content/early/2009/08/14/0900282106.abstract>.

ERICSSON. Ericsson mobility report. 2014.

FICEK, M.; POP, T.; KENCL, L. Active tracking in mobile networks: An in-depth view. Computer Networks, Elsevier, v. 57, n. 9, p. 1936–1954, 2013.

FIFA. Áreas de restrição comercial. Disponível em <http://migre.me/m4MzZ>. Acesso em: 05 de out. 2014. 2013.

FILIPOSKA, S.; MISHKOVSKI, I. Smartphone user's traffic characteristics and modelling. Transactions on Networks and Communications, v. 1, n. 1, 2013. ISSN 2169-6129. Disponível em: <http://scholarpublishing.org/index.php/TNC/article/view/22>.

FRIAS-MARTINEZ, V.; SOGUERO, C.; FRIAS-MARTINEZ, E. Estimation of urban commuting patterns using cellphone network data. In: Proceedings of the ACM SIGKDD International Workshop on Urban Computing. New York, NY, USA: ACM, 2012. (UrbComp '12), p. 9–16. ISBN 978-1-4503-1542-5. Disponível em: <http://doi.acm.org/10.1145/2346496.2346499>.

GAO, L. et al. Quantifying Information Flow During Emergencies. Macmillan Publishers Limited. All rights reserved, v. 4, fev. 2014. Disponível em: <http://dx.doi.org/10.1038/srep03997>.

GILLESPIE, C. S. Fitting heavy tailed distributions: the poweRlaw package. [S.l.], 2014. R package version 0.20.5.

GONZALEZ, M. Unraveling daily human mobility motifs. In: Advances in Social Networks Analysis and Mining (ASONAM), 2013 IEEE/ACM International Conference on. [S.l.: s.n.], 2013. p. xxxviii–xxxviii.

GONZALEZ, M. C.; HIDALGO, C. A.; BARABASI, A.-L. Understanding individual human mobility patterns. Nature, Nature Publishing Group, v. 453, n. 7196, p. 779–782, June 2008.

HALEPOVIC, E.; WILLIAMSON, C. Characterizing and modeling user mobility in a cellular data network. In: Proceedings of the 2Nd ACM International Workshop on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks. New York, NY, USA: ACM, 2005. (PE-WASUN '05), p. 71–78. ISBN 1-59593-182-1. Disponível em: <http://doi.acm.org/10.1145/1089803.1089969>.

HASAN, S. et al. Spatiotemporal patterns of urban human mobility. Journal of Statistical Physics, Springer US, v. 151, n. 1-2, p. 304–318, 2013. ISSN 0022-4715. Disponível em: <http://dx.doi.org/10.1007/s10955-012-0645-0>.

HIDALGO, C. A.; RODRIGUEZ-SICKERT, C. The dynamics of a mobile phone network. Physica A: Statistical Mechanics and its Applications, v. 387, n. 12, p. 3017 – 3024, 2008. ISSN 0378-4371. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0378437108000976>.

HOHWALD, H.; FRíAS-MARTíNEZ, E.; OLIVER, N. User modeling for telecommunication applications: Experiences and practical implications. In: BRA, P. D.; KOBSA, A.; CHIN, D. (Ed.). User Modeling, Adaptation, and Personalization. [S.l.]: Springer Berlin Heidelberg, 2010, (Lecture Notes in Computer Science, v. 6075). p. 327–338. ISBN 978-3-642-13469-2.

ISAACMAN, S. et al. Identifying important places in people's lives from cellular network data. In: Proceedings of the 9th International Conference on Pervasive Computing. Berlin, Heidelberg: Springer-Verlag, 2011. (Pervasive'11), p. 133–151. ISBN 978-3-642-21725-8. Disponível em: <http://dl.acm.org/citation.cfm?id=2021975.2021988>.

ISAACMAN, S. et al. Human mobility modeling at metropolitan scales. In: Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services. New York, NY, USA: ACM, 2012. (MobiSys '12), p. 239–252. ISBN 978-1-4503-1301-8. Disponível em: <http://doi.acm.org/10.1145/2307636.2307659>.

JAIN, R. The art of computer systems performance analysis - techniques for experimental design, measurement, simulation, and modeling. [S.l.]: Wiley, 1991. I-XXVII, 1-685 p. (Wiley professional computing). ISBN 978-0-471-50336-1.

JIANG, B.; JIA, T. Exploring human mobility patterns based on location information of US flights. CoRR, abs/1104.4578, 2011. Disponível em: <http://arxiv.org/abs/1104.4578>.

JIANG, S. et al. A review of urban computing for mobile phone traces: Current methods, challenges and opportunities. In: Proceedings of the 2Nd ACM SIGKDD International Workshop on Urban Computing. New York, NY, USA: ACM, 2013. (UrbComp '13), p. 2:1–2:9. ISBN 978-1-4503-2331-4.

KARAMSHUK, D. et al. Human mobility models for opportunistic networks. Communications Magazine, IEEE, v. 49, n. 12, p. 157–165, December 2011. ISSN 0163-6804.

LABRINIDIS, A.; JAGADISH, H. V. Challenges and opportunities with big data. Proc. VLDB Endow., VLDB Endowment, v. 5, n. 12, p. 2032–2033, ago. 2012. ISSN 2150-8097. Disponível em: <http://dx.doi.org/10.14778/2367502.2367572>.

LANE, N. et al. A survey of mobile phone sensing. Communications Magazine, IEEE, v. 48, n. 9, p. 140–150, Sept 2010. ISSN 0163-6804.

LAURILA, J. K. et al. The mobile data challenge: Big data for mobile computing research. In: . Newcastle, UK: [s.n.], 2012.

LOUAIL, T. et al. From mobile phone data to the spatial structure of cities. Scientific Reports, v. 4, jun. 2014. ISSN 2045-2322. Disponível em: <http://dx.doi.org/10.1038/srep05276>.

MAPS, G. Google maps javascript api v3. Disponível em <https://developers.google.com/maps/documentation/javascript/>. Acesso em: 05 de out. 2014. 2012.

MITCHELL, T. M. Mining our reality. Science, v. 326, n. 5960, p. 1644–1645, 2009. Disponível em: <http://www.sciencemag.org/content/326/5960/1644.short>.

MORLOT, F.; ELAYOUBI, S.; BACCELLI, F. An interaction-based mobility model for dynamic hot spot analysis. In: INFOCOM, 2010 Proceedings IEEE. [S.l.: s.n.], 2010. p. 1–9. ISSN 0743-166X.

MUNJAL, A.; CAMP, T.; NAVIDI, W. C. Smooth: A simple way to model human mobility. In: Proceedings of the 14th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems. New York, NY, USA: ACM, 2011. (MSWiM '11), p. 351–360. ISBN 978-1-4503-0898-4. Disponível em: <http://doi.acm.org/10.1145/2068897.2068957>.

NIR SOFER. Siteshoter v1.42. Disponível em <http://www.nirsoft.net/utils/web_site_screenshot.html>. Acesso em: 05 de out. 2014. 2011.

PULSELLI, R. et al. Computing urban mobile landscapes through monitoring population density based on cell-phone chatting. Design & Nature and Ecodynamics, v. 3, n. 2, p. 121 –134, 2008.

SONG, C. et al. Modelling the scaling properties of human mobility. Nature Physics, Nature Publishing Group, v. 6, n. 10, p. 818–823, set. 2010. ISSN 1745-2473. Disponível em: <http://dx.doi.org/10.1038/nphys1760>.

SONG, C. et al. Limits of predictability in human mobility. Science, v. 327, n. 5968, p. 1018–1021, 2010.

SOPER, D. Is human mobility tracking a good idea? Communications of the ACM, ACM, New York, NY, USA, v. 55, n. 4, p. 35–37, abr. 2012. ISSN 0001-0782. Disponível em: <http://doi.acm.org/10.1145/2133806.2133819>.

TANAHASHI, Y. et al. Inferring human mobility patterns from anonymized mobile communication usage. In: Proceedings of the 10th International Conference on Advances in Mobile Computing & Multimedia. New York, NY, USA: ACM, 2012. (MoMM '12), p. 151–160. ISBN 978-1-4503-1307-0. Disponível em: <http://doi.acm.org/10.1145/2428955.2428988>.

TOOLE, J. L. et al. Inferring land use from mobile phone activity. ACM, New York, NY, USA, p. 1–8, 2012. Disponível em: <http://doi.acm.org/10.1145/2346496.2346498>.

WANG, H. et al. Transportation mode inference from anonymized and aggregated mobile phone call detail records. In: Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on. [S.l.: s.n.], 2010. p. 318–323. ISSN 2153-0009.

WANG, P. et al. Understanding the spreading patterns of mobile phone viruses. Science, v. 324, n. 5930, p. 1071–1076, 2009. Disponível em: <http://www.sciencemag.org/content/324/5930/1071.abstract>.

XAVIER, F. H. Z. et al. Analyzing the workload dynamics of a mobile phone network in large scale events. In: Proceedings of the First Workshop on Urban Networking. New York, NY, USA: ACM, 2012. (UrbaNe '12), p. 37–42. ISBN 978-1-4503-1781-8. Disponível em: <http://doi.acm.org/10.1145/2413236.2413245>.

YAN CHEN ZHAO, Y. F. Z. D. W.-X. W. X.-Y. Universal predictability of mobility patterns in cities. Journal of The Royal Society Interface, n. 100, p. 20140834, 2014. Disponível em: <http://rsif.royalsocietypublishing.org/content/11/100/20140834.abstract>.

YUAN, J.; ZHENG, Y.; XIE, X. Discovering regions of different functions in a city using human mobility and pois. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA: ACM, 2012. (KDD '12), p. 186–194. ISBN 978-1-4503-1462-6. Disponível em: <http://doi.acm.org/10.1145/2339530.2339561>.

8 LISTA DE EVENTOS COLETADOS

Tabela 29 – Coletas realizadas durante o trabalho

Evento	Tipo Evento	Local	Data	Público
Cruzeiro x Grêmio	Futebol – Campeonato Brasileiro	Arena do Jacaré (MG)	06-07-2011	
Cruzeiro x Atlético	Futebol – Campeonato Brasileiro	Arena do Jacaré (MG)	27-08-2011	
Flamengo x Vasco 28	Futebol – Campeonato Brasileiro	Engenhão	28-08-2011	33.206
Arena do Jacaré sem jogo		Arena do Jacaré (MG)	06-11-2011	
Flamengo x Vasco 04	Futebol – Campeonato Brasileiro	Engenhão	04-12-2011	40.004
Engenhão Sem Evento		Engenhão	11-12-2011	
Arena do Jacaré sem jogo		Arena do Jacaré (MG)	18-12-2011	
Réveillon BH	Festa	Praça da Estação	31-12-2011	100.000
Réveillon Recife	Festa	Cabanga	31-12-2011	10.000
Réveillon RJ	Festa	Copacabana	31-12-2011	2,5 Milhões
Réveillon Salvador	Festa	Praia do Forte	31-12-2011	1 Milhão
Réveillon Sp	Festa	Av. Paulista	31-12-2011	2 Milhões
Réveillon BH Sem Evento		Praça da Estação	03-01-2012	
Réveillon Recife Sem Evento		Cabanga	03-01-2012	
Réveiilon RJ Sem Evento		Copacabana	03-01-2012	
Réveillon Salvador Sem Evento		Praia do Forte	03-01-2012	
Réveillon SP Sem Evento		Av. Paulista	03-01-2012	
Show Roger Waters	Show	Engenhão	29-03-2012	50.000
Corinthians x Boca	Futebol – Copa Libertadores	Pacaembu	04-07-2012	
Flamengo x Fluminense	Futebol – Campeonato Brasileiro	Engenhão	08-07-2012	32.591
Amistoso Brasil x África do sul	Futebol – Amistoso	Morumbi	07-09-2012	
Atlético x Fluminense	Futebol – Campeonato Brasileiro	Independência	21-10-2012	20.096
Formula 1	Corrida	Interlagos	25-11-2012	69.984
Cruzeiro x Atlético	Futebol – Campeonato Mineiro	Mineirão	03-02-2013	52.989
Independência sem jogo		Independência	10-03-2013	
México x Itália	Futebol – Copa das Confederações	Maracanã	16-03-2013	
Interlagos sem evento		Interlagos	24-03-2013	
Morumbi Sem evento		Morumbi	07-04-2013	
Brasil x Chile	Futebol – Amistoso	Mineirão	24-04-2013	53.331
Paul McCartney	Show	Mineirão	04-05-2013	53.000
Espanha x Taiti	Futebol – Copa das Confederações	Maracanã	20-06-2013	
Japão x México	Futebol – Copa das Confederações	Mineirão	22-06-2013	
Brasil x Uruguai	Futebol – Copa das Confederações	Mineirão	26-06-2013	
Brasil x Espanha	Futebol – Copa das Confederações	Maracanã	30-06-2013	73.531
Itália x Uruguai	Futebol – Copa das Confederações	Arena das Dunas (RN)	30-06-2013	
Beyonce	Show	Mineirão	11-09-2013	30.000
Mineirão sem jogo		Mineirão	03-10-2013	
Cruzeiro x Grêmio	Futebol – Campeonato Brasileiro	Mineirão	10-11-2013	
Flamengo x Atlético PR	Futebol – Copa do Brasil	Maracanã	27-11-2013	68.857
Colômbia x Uruguai	Futebol – Copa do Mundo	Maracanã	28-06-2014	73.804
Holanda x México	Futebol – Copa do Mundo	Castelão	29-06-2014	58.817
Grécia x Costa Rica	Futebol – Copa do Mundo	Arena Pernambuco	29-06-2014	41.242
Alemanha x Argentina	Futebol – Copa do Mundo	Maracanã	13-07-2014	74.738