Análise da Mobilidade Humana em Eventos de Larga Escala baseada em Chamadas de Telefones Celulares

Faber Henrique Z. Xavier[1], Lucas M. Silveira[1], Jussara M. Almeida[2], Artur Ziviani[3], Carlos Henrique Malab[4], Humberto T. Marques-Neto[1]

[1] Pontifícia Universidade Católica de Minas Gerais (PUC Minas)
30.535-901 - Belo Horizonte - Brasil

[2]Universidade Federal de Minas Gerais (UFMG)
31.270-010 - Belo Horizonte - Brasil

[3]Laboratório Nacional de Computação Científica (LNCC/MCTI)
25.651-075 - Petrópolis - Brasil

[4]Oi Telecom
20.230-070 - Rio de Janeiro - Brasil

{faber.xavier,lmsilveira}@sga.pucminas.br, jussara@dcc.ufmg.br

ziviani@lncc.br, malab@oi.net.br, humberto@pucminas.br

Abstract. *The analysis of the usage of mobile phone networks can help cellphone carriers to evolve and improve their services, as well as foster the well-being of our society with more "intelligent" services based on human mobility patterns. This work analyzes human mobility during different types of large-scale events, specifically soccer matches and New Year's celebrations, using a methodology recently proposed by the authors. Human mobility is analyzed in terms of number of cellphone calls, their inter-arrival and inter-departure times, and their durations during these events. We also make use of heatmaps to analyze the displacement patterns of participants around the time and geographical area of the event. Our results could be used to improve the understanding of human mobility in urban areas during large-scale events.*

Resumo. *A análise dos dados de uso dos recursos das redes de telefonia celular contribui para a melhoria da prestação deste tipo de serviço essencial na sociedade contemporânea, além de poder promover o bem-estar dessa sociedade com serviços cada vez mais "inteligentes" baseados em padrões de mobilidade humana. Este artigo analisa a mobilidade humana durante alguns tipos de eventos de larga escala, especificamente jogos de futebol e Reveillon utilizando uma metodologia recentemente proposta pelos autores. A mobilidade humana é analisada sob o ponto de vista da quantidade, tempos entre início e término, e duração das chamadas de celulares realizadas durante os eventos estudados. Utiliza-se ainda mapas de calor para analisar os padrões de deslocamento dos participantes em torno do horário e na região dos diferentes eventos. Espera-se que os resultados produzidos possam melhorar a compreensão da mobilidade humana em áreas urbanas em decorrência de grandes eventos.*

1. Introdução

A análise de bases de dados que armazenam informações das chamadas de telefones celulares realizadas na rede das operadoras tem se mostrado um grande desafio, sobretudo quando se pretende compreender os padrões de mobilidade de usuários em áreas urbanas muito populosas. Contudo, tal análise pode proporcionar benefícios tanto para as operadoras de telefonia celular quanto para os governos, já que provêm subsídios para melhorias nos serviços prestados à população [González et al. 2008, Soper 2012] em cidades que precisam ser cada vez mais "inteligentes". Os dados das chamadas de voz em uma rede de telefonia celular quando relacionados com o posicionamento geográfico das antenas utilizadas pelos usuários da operadora podem viabilizar a criação de ferramentas para predição da mobilidade humana e também para o planejamento da utilização dos recursos das operadoras [Candia et al. 2008].

O estudo de [Song et al. 2010] mostra que a mobilidade humana pode ser prevista, com razoável precisão, a partir da análise das rotinas diárias das pessoas. A previsão da movimentação dos usuários é importante para auxiliar as operadoras de telefonia celular a definir a infraestrutura necessária para garantir o uso dos seus serviços. Contudo, as operadoras normalmente encontram dificuldades para planejar o uso de sua rede em situações que geram uma carga de trabalho fora do padrão, como durante a realização de grandes eventos [Bleicher 2012]. Como exemplos desses grandes eventos, podemos mencionar a final de um campeonato de futebol, um grande show de rock, a véspera de Ano Novo e a realização dos Jogos Olímpicos. Mesmo se forem conhecidos o público esperado, o horário e a data, 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 et al. 2011].

Especificamente, como é de amplo conhecimento, a Copa do Mundo de Futebol de 2014 e os Jogos Olímpicos de 2016 serão realizados no Brasil. Esses eventos esportivos de larga escala atraem um grande público, o qual se move em torno de uma determinada área para assistir as principais competições ou eventos satélites menores, bem como para realizar passeios turísticos e conhecer as atrações do país sede. Espera-se que as áreas urbanas das cidades sede desses eventos recebam até centenas de milhares de participantes, com seus dispositivos móveis, contribuindo assim para um aumento considerável da carga imposta na rede de telefonia celular. Portanto, os padrões de comportamento de mobilidade humana, inferidos a partir do histórico de uso de serviços de telefonia móvel, podem apoiar o desenvolvimento de melhores estratégias de gestão de recursos e serviços de comunicação tanto para as pessoas que frequentam esses eventos de grande porte quanto para aquelas que vivem nessas cidades.

Este trabalho analisa características da mobilidade humana durante alguns tipos de eventos de larga escala, especificamente jogos de futebol e Réveillon, a partir de bases de dados com informações totalmente anonimizadas sobre chamadas de voz realizadas na rede de telefonia celular de uma grande operadora brasileira. Para tal, foi utilizada uma metodologia de análise previamente proposta pelos autores [Xavier et al. 2012], aplicada a bases de dados contendo as chamadas feitas durante dois jogos de futebol no Rio de Janeiro e durante a comemoração do Ano Novo em Belo Horizonte, Recife e Salvador. A mobilidade humana é analisada em termos da quantidade, tempos entre início e término, e duração das chamadas feitas em uma região e período próximos ao local e horário de cada

evento de larga escala considerado, bem como em um dia comum (sem evento) na mesma região e horário. Utiliza-se ainda mapas de calor para analisar os padrões de deslocamento dos participantes em torno do horário e região de realização do evento.

Este artigo estende os resultados preliminares relatados em [Xavier et al. 2012], focados no impacto da dinâmica da carga de trabalho na rede de telefonia celular durante apenas um único tipo específico de evento de larga escala (jogo de futebol). Em particular, ele generaliza e valida a metodologia proposta em [Xavier et al. 2012] para outros tipos de eventos (comemoração de Ano Novo) com o propósito de compreender a mobilidade humana em eventos de larga escala que acontecem, normalmente, em cidades populosas. Este artigo também estende as análises feitas anteriormente para incluir os tempos entre inícios e entre términos de chamadas sucessivas e as durações de chamadas, aspectos importantes relacionados à carga imposta nas operadoras que não foram analisados em [Xavier et al. 2012]. Espera-se que os resultados produzidos possam melhorar o entendimento sobre a mobilidade humana em decorrência de grandes eventos, provendo subsídios para o melhor gerenciamento e planejamento das redes das operadoras e para o desenvolvimento de novas aplicações que tornem as cidades mais "inteligentes".

O restante deste artigo está organizado como segue. A Seção 2 discute trabalhos relacionados. A metodologia de análise adotada e as bases de dados utilizadas são apresentadas nas Seções 3 e 4, respectivamente. Os principais resultados da análise realizada são discutidos na Seção 5. Conclusões e trabalhos futuros são apresentados na Seção 6.

2. Trabalhos Relacionados

O crescimento das redes de telefonia móvel para atender a demanda de comunicação atual requer estudos para melhor compreender o comportamento de seus usuários. Os padrões de movimentação humana (de onde as pessoas saíram e para onde elas foram), em diferentes períodos do dia, podem ser analisados a partir dos registros de chamadas feitas pelos usuários de telefonia celular, mantidos pelas operadoras [Song et al. 2010].

Para melhor compreender a carga de trabalho gerada em uma rede móvel em decorrência de um evento fora da rotina de um determinado ambiente é importante realizar análises do uso desta rede tanto em dias com esse evento quanto em um dia típico da rotina do ambiente. Assim, é possível criar modelos matemáticos que ajudem a compreender o comportamento dessa carga tanto em eventos de grande escala quanto em situações corriqueiras [Song et al. 2010]. A análise de um único ambiente (e.g., uma cidade) pode facilitar o entendimento da movimentação das pessoas dentro deste espaço e assim identificar os tipos de usuários que participam dos eventos daquela região [Calabrese et al. 2010]. Adicionalmente, percebe-se que o tipo de carga de trabalho gerado em cada tipo evento tem características específicas, sendo possível até mesmo identificar a sua natureza (show de rock, acidente, etc) a partir da análise dos dados dessa carga de trabalho [Bagrow et al. 2011].

Uma forma de analisar o comportamento humano é através da modelagem da sua mobilidade. Em [González et al. 2008] é utilizado um modelo matemático de predição de mobilidade que utiliza os dados gerados pela telefonia móvel para identificar os padrões de movimentação com rotinas diárias. Esses grupos de pessoas podem ser divididos em categorias de acordo com o tipo de dado [Toole et al. 2012] ou através de focos de ligações na região analisada [Morlot et al. 2010].

A predição da mobilidade humana pode proporcionar benefícios para a sociedade [González et al. 2008]. Por exemplo, ela provê informações para realização de planejamento urbano visando melhorar o tráfego de veículos, apoiar a tomada de decisões em situações de emergência e até mesmo melhorar a gestão de saúde pública para controle da disseminação de uma doença [Balcan et al. 2009]. Mais ainda, ela pode fornecer subsídios para ajudar lojistas e centros comerciais a oferecerem produtos e serviços que levem em conta a maneira como as pessoas se movimentam em uma determinada região em diferentes períodos [Song et al. 2010].

Este artigo analisa diferentes tipos de eventos como feito em [Batty et al. 2003] e, como em [Song et al. 2010], explora uma metodologia para analisar a mobilidade humana a partir de dados reais de chamadas realizadas por usuários anônimos de uma operadora de telefonia celular. Entretanto, enquanto em [Song et al. 2010] os autores analisam chamadas feitas durante todo o dia, nós aqui dividimos o período de análise em intervalos relacionados ao evento considerado e focamos na mobilidade de pessoas direta ou indiretamente envolvidas com ele (p.ex: participantes).

3. Metodologia de Caracterização

3.1. Tratamento de Dados

Em [Xavier et al. 2012], os autores propuseram uma metodologia de análise da dinâmica da carga de trabalho de uma rede de telefonia celular durante grandes eventos, focando especificamente em jogos de futebol. A metodologia foi projetada visando prover subsídios para o gerenciamento e o planejamento de capacidade da infraestrutura de rede da operadora para atender as demandas durante eventos de larga escala. Ela utiliza como entrada chamadas de telefones celulares, com usuários anonimizados mas unicamente identificados, explicitando as antenas (i.e., suas localizações geográficas) da operadora utilizadas quando do início e término de cada chamada. Uma mesma chamada associada a duas antenas diferentes implica em deslocamento do usuário durante a chamada.

A metodologia de análise apresentada em [Xavier et al. 2012] foi elaborada com o propósito de identificar os usuários de uma operadora de telefonia celular envolvidos direta ou indiretamente com o evento de larga escala (ou seja, participantes do evento), e analisar a carga resultante das ligações feitas por eles em período próximo ao evento. A premissa principal é que usuários presentes próximos ao local do evento, durante sua realização, estavam envolvidos com o evento. Assim, objetiva-se identificar os usuários que estavam na região do evento durante o mesmo, de onde vieram e para onde foram após o seu término, a partir de informações anonimizadas das ligações realizadas por esses usuários utilizando os recursos das antenas que cobrem tanto o local do evento quanto as suas principais vias de acesso. A partir da identificação destes usuários e de suas chamadas, pode-se analisar a carga imposta por eles nessas antenas.

A metodologia consiste primeiramente em definir uma região geográfica e uma janela de tempo que delimitam, espacial e temporalmente, a realização do evento. Essa definição deve 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 antenas que cobrem essa região e cujas cargas poderão ser afetadas pelo público do evento. Assim, as análises deverão focar nas chamadas realizadas nestas antenas. Além disto, é preciso focar nas chamadas feitas durante um período em torno do horário do evento. Para

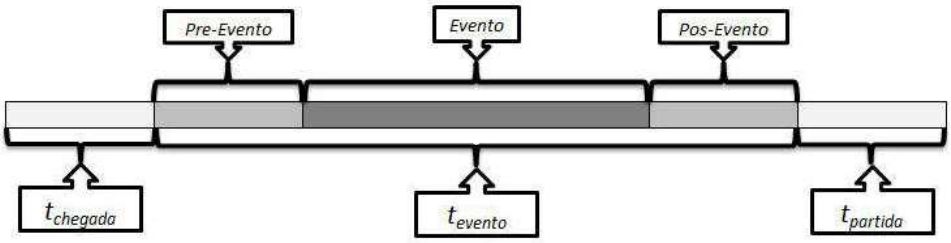


Figura 1. Linha do tempo para a análise da carga durante um evento.

tal, adota-se a notação da linha de tempo mostrada na Figura 1. O período marcado como *Evento* corresponde ao período de realização do evento em si.

O segundo passo da metodologia consiste em identificar os usuários que fizeram pelo menos uma chamada em uma das antenas selecionadas em horário próximo ao evento. Note que é possível que os usuários não façam muitas chamadas durante a realização do evento, dado seu envolvimento com ele. Isto é esperado principalmente para alguns tipos de evento, como jogos de futebol. A fim de identificar os usuários envolvidos com o evento, aqui chamados de *espectadores*, é preciso estender o período analisado. Para tal, são definidos os períodos *Pre-Evento* e *Pos-Evento*, que antecede e sucede, respectivamente, o evento. O período total desde o início do *Pre-Evento* até o término do *Pos-Evento* é chamado de t_{evento}. Usuários que fizeram pelo menos uma chamada durante t_{evento} são considerados *espectadores*, i.e., são alvo das análises realizadas.

Note que o período *Pre-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 que o *Pos-Evento* deve capturar um período durante o qual as pessoas 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.

Por fim, identifica-se o subconjunto de espectadores 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 estes usuários, são definidos dois outros períodos — $t_{chegada}$ e $t_{partida}$ — durante os quais pessoas estão se movimentando para a região do evento ou partindo dela para outro local. Utilizando a localização geográfica das antenas onde cada chamada iniciou e terminou, pode-se determinar de onde estes espectadores vieram e para onde foram, após o fim do evento. Ou seja, é possível estudar a *dinâmica da carga* nas antenas localizadas nas principais rotas de acesso à região do evento. No contexto de eventos de grande porte como a Copa do Mundo e os Jogos Olímpicos, estas rotas possivelmente conectam estádios com regiões de hotéis, restaurantes e aeroportos. A identificação dessas rotas pode ser explorada para diversos serviços direcionados, tais como propaganda ou orientações ao público.

3.2. Análise das Cargas de Trabalho

A análise da carga imposta nas antenas da região do evento pelos usuários espectadores, identificados conforme descrito na Seção 3.1, foca nos seguintes aspectos: número de chamadas, tempos entre inícios de chamadas sucessivas em uma mesma antena, tempos entre términos de chamadas sucessivas e duração das chamadas. A análise dos tempos entre inícios e términos de chamadas sucessivas é de grande valor para investigar como a

vazão das chamadas se compara ao fluxo de entrada. Por sua vez, esta análise é importante para as operadoras, pois uma sobrecarga em suas antenas traz perda na qualidade de serviço prestado e um possível bloqueio de chamadas dos usuários, podendo refletir em perda financeira e marketing negativo. Da mesma forma, a análise da duração das chamadas é importante para entender a ocupação das antenas durante o período considerado. As análises dos tempos entre inícios, entre términos e duração das chamadas são contribuições deste artigo em relação a [Xavier et al. 2012], que não aborda estes aspectos.

Os aspectos acima são aqui analisados não somente de forma agregada, considerando o período t_{evento} inteiro, mas também separadamente para intervalos de 15 minutos sucessivos durante aquele período. Assim, pode-se avaliar a variação da carga durante estes vários intervalos. Para avaliar o impacto do evento de larga escala na carga das antenas da região, a carga durante o intervalo t_{evento} é comparada com a carga na mesma região (mesmas antenas) em período equivalente em um dia comum, sem evento.

Para analisar a dinâmica da carga nas antenas localizadas nas rotas de acesso ao local do evento são utilizados mapas de calor para representar a intensidade de utilização de cada antena pelo serviço de telefonia móvel. Os mapas são gerados por uma função da Google Maps Javascript API V3 [Google Maps 2012], que recebe como entrada a latitude, a longitude e o número de ligações de cada antena analisada. A região em que a antena está localizada é apresentada com uma cor, indicando o número de chamadas recebidas por ela durante o intervalo analisado, conforme padrão de cores: quanto mais escura a cor, maior o número de chamadas recebido.

O resultado dessa função é uma imagem do mapa da região, com a sobreposição do mapa de calor indicando a localização das antenas e o volume de ligações em cada uma, conforme padrão de cores. Assim, é possível identificar as principais rotas de acesso ao local do evento analisando os mapas gerados para os períodos $t_{chegada}$, t_{evento} e $t_{partida}$.

4. Bases de Dados

A metodologia apresentada na Seção 3 foi aplicada a várias bases de dados disponibilizadas por uma grande operadora de telefonia móvel brasileira. Cada base contém todas as chamadas realizadas a partir de telefone celular em uma determinada cidade e em um dado dia. Para cada chamada, a base fornece os seguintes campos:

- identificador único associado à chamada, que não permite a identificação do número do telefone celular do usuário;
- 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;
- tempos de início e fim de cada chamada;
- localizações geográficas (latitude e longitude) das antenas onde a chamada foi iniciada e terminada;
- identificadores dos setores das antenas onde a chamada foi iniciada e terminada[1].

As bases de dados analisadas foram selecionadas considerando dias e cidades em que houve dois tipos de eventos de grande escala: jogos de futebol e Réveillon.

[1]Cada antena é dividida em 3 setores de 120° responsável por um terço da área de cobertura da antena.

Duas bases analisadas são de jogos realizados no Estádio Olímpico João Havelange (Engenhão) nos dias 28 de Agosto e 4 de Dezembro de 2011. Os jogos foram entre Vasco e Flamengo pelo Campeonato Brasileiro de Futebol de 2011. O primeiro jogo, em 28 de Agosto, foi a primeira vez que os dois times se enfrentaram pelo campeonato. Foi reportado que 33.206 pessoas estiveram presentes no Engenhão durante este jogo. Já no segundo jogo, em 4 de Dezembro, a audiência reportada chegou a 34.064 pessoas[2].

Para fins de comparação, também foi analisada uma base de dados com chamadas realizadas em 30 de Outubro de 2011, no Rio de Janeiro. Assim como os dois dias de jogos, este terceiro dia foi um domingo, mas sem nenhum jogo de futebol no estádio Engenhão e em nenhum outro estádio da cidade[3]. Note que a movimentação de pessoas na região do Engenhão é geralmente grande, independentemente do dia, já que ele está localizado em uma área densamente povoada. Apesar disso, acredita-se que a análise dos dados coletados em um dia sem jogo provê um importante ponto de referência para avaliar a dinâmica e o impacto na carga das antenas da operadora devido a eventos de grande escala (um jogo de futebol, no caso).

Também foram analisadas bases de dados contendo as chamadas realizadas durante o Réveillon de 2012 em Belo Horizonte, Recife e Salvador. No caso de Belo Horizonte, considerou-se a festa de Réveillon realizada na Praça da Estação, com um público presente de aproximadamente 100.000 pessoas. Já para Recife, considerou-se a festa para 10.000 pessoas realizada no Clube Cabanga, enquanto que em Salvador, considerou-se o evento realizado na Praia do Farol da Barra, onde compareceram aproximadamente 1.000.000 de pessoas[4]. Mais uma vez, para fins de comparação, também foram analisadas bases de dados das mesmas localidades, coletadas no dia 3 de Janeiro de 2012.

5. Resultados

5.1. Aplicação da Metodologia nas Bases de Dados

Conforme descrito na Seção 3, o primeiro passo da metodologia de análise consiste em definir a região do evento, e consequentemente as antenas da operadora que cobrem esta região, bem como o período de análise. No caso dos jogos de futebol, a região do evento foi definida como sendo a área geográfica cobrindo um raio de 1.500 metros do estádio Engenhão. Neste caso, 4 antenas, localizadas dentro desta área, foram selecionadas para análise. Para as bases com dados do Réveillon de 2011, as regiões do evento foram definidas também partindo de um raio de 1.500 metros do local oficial do evento: a Praça da Estação em Belo Horizonte, o Clube Cabanga em Recife e a Praia do Farol da Barra em Salvador. Considerando esta região, foram selecionadas 3 antenas para Belo Horizonte e Recife; e 6 para Salvador. Optou-se por estender a área de análise além do local oficial de realização do evento a fim de capturar chamadas realizadas nas proximidades do evento, que provavelmente foram feitas por pessoas que participaram dele. Por simplicidade, adotou-se o mesmo raio de 1.500 metros para todas as bases.

Quanto à linha de tempo considerada na análise, a janela total de tempo adotada para os dias de jogos de futebol foi o intervalo entre 14:30 e 19:15. Cada partida é iniciada

[2]Fonte: http://futpedia.globo.com/campeonato/campeonato-brasileiro.

[3]Como o Rio de Janeiro adota o horário de verão, os tempos de início e término de cada chamada foram deslocados uma hora para trás para as bases de Outubro e Dezembro.

[4]A audiência de cada evento foi obtida em http://g1.globo.com e http://www.cabanga.com.br/home/.

Tabela 1. Visão geral das bases de dados de 2011 (dias com evento em negrito).

	Dia	# Chamadas de Usuários	# Chamadas de Espectadores	# Usuários	# Espectadores
Engenhão	**28/08**	**50.138**	**3.362**	**28.921**	**1.366**
	04/12	**47.776**	**4.284**	**27.290**	**1.754**
	31/10	41.021	1.270	25.313	691
Belo Horizonte	**31/12**	**11.348**	**5.187**	**5.729**	**1.938**
	03/01	4.939	779	2.923	365
Recife	**31/12**	**15.189**	**9.951**	**7.116**	**3.566**
	03/01	3.505	924	2.116	444
Salvador	**31/12**	**36.805**	**12.826**	**19.215**	**7.458**
	03/01	6.464	1.019	4.044	689

às 16:00, com uma duração total (*Evento*), incluindo intervalo entre tempos, de 105 minutos. Foi escolhido o valor de 45 minutos para $t_{chegada}$, $t_{partida}$, *Pre-Evento* e *Pos-Evento*. Para os dias de Réveillon, foi analisado o intervalo entre 21:45 e 2:30, assumindo o início do evento às 23:15 e também com duração de 105 minutos. Por simplicidade, foi utilizado também 45 minutos como valor para $t_{chegada}$, $t_{partida}$, *Pre-Evento* e *Pos-Evento*. Note que a mesma região, as mesmas antenas e as mesmas linhas do tempo foram utilizadas para as análises dos dias sem evento em cada localidade.

5.2. Visão Geral das Bases de Dados

A Tabela 1 apresenta uma visão geral das bases de dados analisadas, informando o número total de chamadas e de usuários que fizeram chamadas (colunas 3 e 5) em alguma antena da cidade durante o período entre o início de $t_{chegada}$ até o término de $t_{partida}$. Os números totais de antenas nas bases do Rio de Janeiro (jogos), Belo Horizonte, Recife e Salvador são 146, 38, 30 e 58, respectivamente. A coluna 4 apresenta o número total de chamadas feitas por espectadores nas antenas da região do evento, enquanto que a coluna 6 apresenta o número desses usuários, para o mesmo período. As linhas em negrito correspondem a dias com evento de larga escala na região selecionada.

Fixada a região, nota-se um grande aumento nos números de chamadas e de usuários nos dias com evento, quando comparado aos dias sem evento. Isto vale principalmente para chamadas feitas por espectadores nas antenas da região, com um fator de aumento entre 65% e 137% para os jogos. Para os dias de Réveillon, o número de chamadas por espectadores aumenta por um fator bem maior (entre 5,65 e 47,66), o que reflete o período de análise (noite e madrugada) e o tipo de evento que fomenta uma maior interação entre as pessoas. É interessante notar também o grande aumento no número total de chamadas feitas em todas as antenas da cidade nos dias de Réveillon, o que era esperado já que é um evento celebrado por todos e não restrito às regiões selecionadas. Isto mostra que a preparação, em termos de gerenciamento e planejamento da capacidade da operadora, para eventos que influenciam grande parte da população deve ser diferente daqueles destinados a um grupo específico, como uma partida de futebol.

5.3. Número de Chamadas

Esta seção analisa a distribuição temporal do número de chamadas realizadas por espectadores durante o período t_{evento}, focando nas antenas da região selecionada. A Figura 2 mostra este número para sucessivos intervalos de 15 minutos para as regiões do Engenhão e do Réveillon em Recife. Essa figura também mostra os resultados para os dias sem evento. Os resultados para o Réveillon em Belo Horizonte e Salvador são semelhantes, sendo portanto omitidos; o mesmo ocorre para o jogo do dia 28/08/2011.

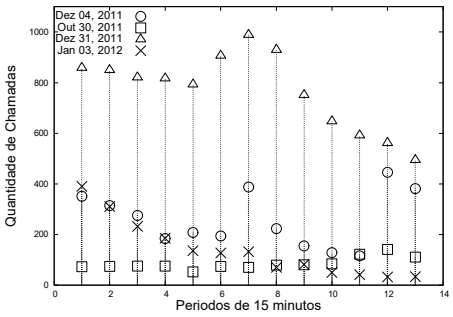


Figura 2. Número de chamadas — Réveillon em Recife e Fla x Vas (04/12/11).

Tabela 2. Tabela de análise de IAT.

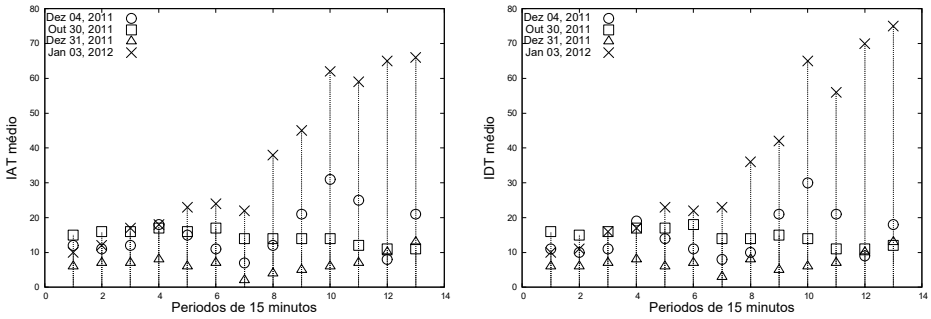
	IAT			IDT		
	IAT médio (s)	S_{inter}	$\overline{S}_{intra}$	IDT médio (s)	S_{inter}	$\overline{S}_{intra}$
Belo Horizonte	7,19	2,64	2,96	7,51	2,59	2,73
Recife	5,65	4,05	1,30	5,69	4,14	1,33
Salvador	12,45	8,02	11,63	8,46	2,47	5,92
Fla x Vas (04/12)	16,13	15,58	8,22	15,44	14,24	7,70
Fla x Vas (28/08)	33,84	46,71	29,82	32,67	44,25	29,69

Como mostrado em [Xavier et al. 2012], para melhor compreender o comportamento da carga durante grandes eventos, a análise desta deve ser feita de forma temporal. Partes do evento onde possa existir um comportamento diferente devem ser consideradas. Durante os eventos analisados, o intervalo da partida e a meia-noite são os períodos onde o comportamento dos usuários sofre alteração. Durante esses períodos, os usuários tendem a realizar um número maior de chamadas. O mesmo comportamento notado em [Xavier et al. 2012] com relação a um volume maior de chamadas no inicio da análise também foi percebido durante a festa de Réveillon. Essas chamadas são provenientes dos usuários que chegam mais cedo no local do evento. Ao contrário do que ocorre para a partida de futebol, onde o número de ligações sofre um rápido aumento após o término do evento, o número de ligações decresce após o término da festa de Ano Novo.

5.4. Intervalo entre chamadas

Os tempos entre inícios e os tempos entre términos de chamadas sucessivas em uma mesma antena, aqui denominados de *InterArrival Times* (IAT) e *InterDeparture Times* (IDT), são úteis para avaliar a carga imposta nas antenas (ver Seção 3.2). A Tabela 2 mostra os valores de IAT e IDT médios calculados para as antenas que cobrem a região do evento durante t_{evento}. Essas antenas foram escolhidas por serem as responsáveis por suportar a maior parte da carga gerada pelo evento. As colunas S_{inter} e $\overline{S}_{intra}$ representam, respectivamente, o desvio padrão dos tempos médios computados para diferentes antenas e a média dos desvios padrões computados para cada antena durante o intervalo t_{evento}.

A Tabela 2 mostra que o comportamento das cargas de trabalho varia com o tipo de evento. Os espectadores do Réveillon iniciam e finalizam mais chamadas em um intervalo de tempo menor do que os torcedores dos jogos de futebol analisados. Esse comportamento pode ser percebido nas Figuras 3(a) e 3(b), que mostram os valores médios de IAT e IDT para intervalos sucessivos de 15 minutos do t_{evento}. Como as curvas de todas as festas de Réveillon, bem como as curvas dos jogos de futebol, são semelhantes entre si, essas figuras mostram apenas os resultados para uma festa de Réveillon (BH) e um jogo (FLAxVAS em 04/12/2011), além dos resultados referentes a dias sem evento nas respectivas regiões de ambos. Isto é necessário em decorrência da restrição de espaço.



(a) IAT (b) IDT

Figura 3. IAT e IDT — Réveillon em BH e Fla x Vas (04/12/2011).

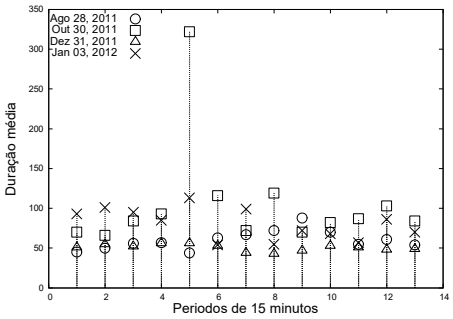


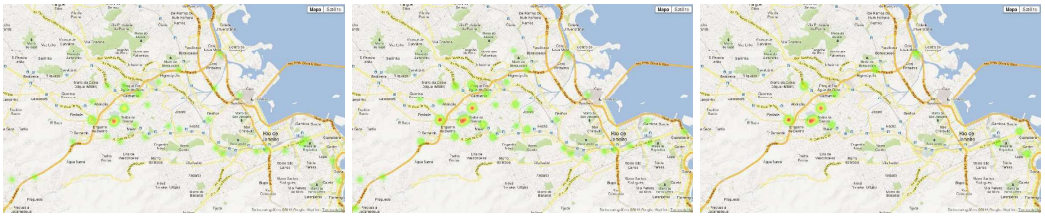
Figura 4. Duração de chamadas — Réveillon em BH e Fla x Vas (28/08/2011).

Como mostrado em [Xavier et al. 2012] os espectadores em uma partida de futebol tendem a realizar mais chamadas durante o pré-jogo, intervalo e pós-jogo. Estes períodos são caracterizados por baixos valores de IAT. Por outro lado, perecebe-se que os usuários que participam de celebrações de Réveillon realizam muito mais chamadas ao longo de todo o t_{evento}. Os valores de IAT de chamadas nas antenas próximas ao Engenhão se aproximam dos mesmos valores nas antenas da região da festa de Réveillon somente em alguns momentos, como por exemplo no intervalo do jogo (período 6). Uma análise das Figuras 3(a) e 3(b) mostra que em alguns intervalos a curva do IAT está mais próxima do eixo X do que a respectiva curva do IDT. Este comportamento indica que nesses intervalos os usuários iniciam mais chamadas do que finalizam. Ou seja, eles, em conjunto, ficam por mais tempo utilizando os recursos da operadora de telefonia celular. Isto pode gerar uma sobrecarga na antena, resultando assim, em uma diminuição da qualidade de serviço percebido pelos usuários.

5.5. Análise da duração das chamadas

Além dos valores de IAT e IDT, a análise das durações das chamadas dos usuários é fundamental para a compreensão do uso das antenas. A Figura 4 mostra a duração média das chamadas durante períodos de 15 minutos no t_{evento}. Chamadas onde a duração ultrapassa o período de 15 minutos são consideradas apenas no intervalo de sua criação. Como na seção anterior, as antenas analisadas são apenas as da região do evento.

Como identificado na Seção 5.4, a redução do IAT, que aponta o início de um número maior de ligações em um determinado intervalo de tempo, e a redução do IDT, que é reflexo do registro de término de um grande número de chamadas, indicam a existência de um grande número de chamadas de curta duração. Esta hipótese sobre a duração das



(a) $t_{chegada}$ (b) t_{evento} (c) $t_{partida}$

Figura 5. Movimentação dos torcedores durante o jogo FLAxVAS em 04/12/2011.

chamadas realizadas ao longo do t_{evento} é confirmada na Figura 4. Esta figura mostra que as chamadas realizadas durante as partidas de futebol ou durante as festas de celebração de Réveillon analisadas, têm curta duração.

Enfim, juntamente com os resultados de [Xavier et al. 2012], podemos concluir que a carga de trabalho dos eventos estudados é formada por um conjunto de chamadas de curta duração realizadas pelos participantes ao longo do evento. Contudo, picos de chamadas ocorrem em momentos distintos relacionados à natureza do evento. Ou seja, torcedores realizam menos chamadas durante a efetiva realização da partida de futebol e no Réveillon os participantes das festas fazem chamadas ao longo de todo t_{evento}.

5.6. Mobilidade dos Espectadores: Mapas de Calor

Nesta seção, mapas de calor são utilizados para analisar a dinâmica da carga devido aos espectadores antes, durante e após o evento, e assim inferir as rotas mais usadas para acesso à região do evento. As Figuras 5(a) a 5(c) mostram esses mapas para o Rio de Janeiro no dia 04/12/2011 antes ($t_{chegada}$), durante (t_{evento}) e após ($t_{partida}$) o jogo de futebol. Por restrições de espaço, apenas os resultados para este dia são mostrados, embora padrões de mobilidade semelhantes tenham sido obtidos para os demais eventos.

Nesses mapas é possível perceber três rotas para chegar à região do evento. Uma rota vinda do sul, saindo da Barra da Tijuca, outra contornando pelo norte do estádio e a terceira vindo do leste. Esta última rota é a linha de metrô. Devido à preferência pelo transporte público, provavelmente para evitar o trânsito de veículos na chegada e saída do evento, é a rota mais utilizada. Embora os mapas correspondentes ao dia sem evento não tenham sido mostrados, ressalta-se que em comparação com eles, os mapas das Figuras 5(a) a 5(c) mostram um aumento do número de chamadas em torno do horário e nas proximidades do evento, confirmando a influência do evento na carga imposta às antenas localizadas na região do evento e nas três principais rotas de acesso a ele.

6. Conclusões e Trabalhos Futuros

Com a análise de chamadas de celular realizadas durante alguns jogos do Campeonato Brasileiro de Futebol de 2011 e também durante algumas comemorações de Réveillon de 2011-2012, foi possível identificar padrões de comportamento de usuários de telefonia móvel quando participam de grandes eventos. Observa-se que esse comportamento depende da natureza do evento. Este trabalho apresentou análises do IAT, do IDT e da duração de chamadas em uma metodologia previamente proposta pelos autores. Espera-se que com essas análises, as operadoras de telefonia móvel possam adequar a sua rede para fornecer melhores serviços, principalmente, durante a Copa do Mundo e as Olimpíadas que serão sediadas no Brasil nos próximos anos.

Como trabalhos futuros, pretende-se expandir a análise e aplicação da metodologia para novas bases de dados e diferentes tipos de eventos como shows, corridas automobilísticas, entre outros. Busca-se assim melhor compreender o comportamento dos usuários em eventos de grande escala e propor soluções para um serviço de maior qualidade. Pretende-se também analisar o envio de SMS e a utilização da rede 3G/4G, visando adaptar a metodologia proposta para comparar o comportamento dos usuários para diferentes tipos de eventos e diferentes tecnologias de comunicação móvel. Outras direções possíveis para trabalhos futuros contemplam a exploração dos resultados destas análises no desenvolvimento de aplicativos para aparelhos móveis que tenham como foco eventos de larga escala. Além disto, espera-se aprimorar a ferramenta para geração automática e gerenciamento dos mapas de calor, com o objetivo de oferecer uma melhor análise e visualização da dinâmica da carga em redes celulares devido a grandes eventos.

Referências

Bagrow, J. P., Wang, D., e Barabási, A.-L. (2011). Collective response of human populations to large-scale emergencies. *PLoS ONE*, 6(3):e17680.

Balcan, D., Colizza, V., Gonçalves, B., Hu, H., Ramasco, J. J., e Vespignani, A. (2009). Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*, 106(51):21484–21489.

Batty, M., Desyllas, J., e Duxbury, E. (2003). The discrete dynamics of small-scale spatial events: agent-based models of mobility in carnivals and street parades. *International Journal of Geographical Information Science*, 17(7):673–697.

Bleicher, A. (2012). The on-demand olympics. *IEEE Spectrum*, 49(7):9–10.

Calabrese, F., Pereira, F., Di Lorenzo, G., Liu, L., e Ratti, C. (2010). The geography of taste: Analyzing cell-phone mobility and social events. Em Floréen, P., Krüger, A., e Spasojevic, M., editors, *Pervasive Computing*, volume 6030 of *Lecture Notes in Computer Science*, pág. 22–37. Springer Berlin/Heidelberg.

Candia, J., González, M. C., Wang, P., Schoenharl, T., Madey, G., e Barabási, A.-L. (2008). Uncovering individual and collective human dynamics from mobile phone records. *Journal of Physics A: Mathematical and Theoretical*, 41(22):224015.

González, M. C., Hidalgo, C. A., e Barabási, A.-L. (2008). Understanding individual human mobility patterns. *Nature*, 453:779–782.

Google Maps (2012). Google maps javascript api v3.

Morlot, F., Elayoubi, S. E., e Baccelli, F. (2010). An interaction-based mobility model for dynamic hot spot analysis. Em *Proceedings of the 29th conference on Information communications*, INFOCOM'10, pág. 2294–2302, Piscataway, NJ, USA. IEEE Press.

Song, C., Qu, Z., Blumm, N., e Barabási, A.-L. (2010). Limits of predictability in human mobility. *Science*, 327(5968):1018–1021.

Soper, D. (2012). Is human mobility tracking a good idea? *Communications of the ACM*, 55(4):35–37.

Toole, J. L., Ulm, M., González, M. C., e Bauer, D. (2012). Inferring land use from mobile phone activity. Em *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*, UrbComp '12, pág. 1–8, New York, NY, USA. ACM.

Xavier, F. H. Z., Silveira, L. M., Almeida, J. M., Ziviani, A., Malab, C. H. S., e Marques-Neto, H. (2012). Analyzing the workload dynamics of a mobile phone network in large scale events. Em *Proceedings of the UrbaNe Workshop – ACM CoNEXT 2012*.