

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Especialização em Matemática para Professores

**VISUALIZAÇÃO DE DADOS DE UMA REDE DE
TELEFONIA CELULAR UTILIZANDO FIGURAS
GEOMÉTRICAS PROJETADAS NO GOOGLE MAPS**

Samuel Contin Monteiro Gomes

Belo Horizonte
2011

Samuel Contin Monteiro Gomes

**Visualização de dados de uma rede de telefonia
celular utilizando figuras geométricas projetadas no
Google Maps**

Monografia apresentada ao Programa de Pós-Graduação em Matemática para Professores do Ensino Básico da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Especialista em Matemática.

Orientador: Marcelo Azevedo Costa

**Belo Horizonte
2011**

RESUMO

Com a evolução da tecnologia e a popularização dos aparelhos celulares surgiu a necessidade de coletar, visualizar e analisar os dados estatísticos fornecidos pelos equipamentos utilizados nas redes de telefonia celular. Este trabalho descreve um método de visualização de dados estatísticos, georeferenciados. Os dados trabalhados foram coletados de uma rede celular e se referem ao tráfego telefônico cursado nos mais diferentes nós desta rede. Para atingir este objetivo foi utilizado o software gratuito R que permite a visualização dos dados em mapas da plataforma Google Maps. Com a análise destes dados podemos realimentar o sistema e planejar o investimento feito pelas empresas, proporcionando uma maior eficiência e retorno do investimento.

SUMÁRIO

1	INTRODUÇÃO.....	6
2	MOTIVAÇÃO	8
3	METODOLOGIA.....	11
4	ESTATÍSTICAS DESCRITIVAS.....	13
5	GEOREFERENCIAMENTO	17
6	RESULTADOS OBTIDOS	23
7	CONCLUSÃO	31
8	BIBLIOGRAFIA.....	32

LISTA DE FIGURAS

Figura 1 – A chamada telefônica. Fonte: Teleco.com.br	8
Figura 2 - Topologia de rede celular GSM	8
Figura 3 - Torre celular e suas antenas diretivas	10
Figura 4 – Exemplo de Box Plot de uma série temporal, extratificado por dia da semana. ..	14
Figura 5 - Exemplo de Histograma.....	15
Figura 6 – Diagrama de dispersão	16
Figura 7 – Gráfico de linha.....	16
Figura 8 – Geóide: Modelo físico da superfície terrestre (fonte: UFRGS, 2009).....	18
Figura 9 – Elipsóide de revolução	18
Figura 10 – Diferentes opções de visualização do mapa, utilizando RGoogleMaps	21
Figura 11 – Células e os respectivos setores, mapa estilo <i>Terrain</i>	23
Figura 12 – Histograma dos dados trabalhados.....	24
Figura 13 – Gráfico de linha das séries envolvidas	25
Figura 14 – Gráfico Box Plot setor GVTN13	26

1 INTRODUÇÃO

Desde o início da civilização humana o homem vem buscando formas e métodos para visualização de dados. A origem da visualização está nos diagramas geométricos, nas tabelas de posições de estrelas e nos mapas. O primeiro registro que se tem é um mapa escavado em pedra, de uma cidade da região da Babilônia, datado da época de 6200 a.C. Em 950 d.C. temos o primeiro registro de uma figura utilizando variáveis graficamente (posição do sol e da lua ao longo do ano). A partir do século 16 novas técnicas e instrumentos foram sendo desenvolvidos e com eles formas mais precisas de representação do conhecimento. Mas foi na primeira metade do século XIX que tivemos uma explosão no crescimento dos gráficos estatísticos e do mapeamento temático. A grande maioria dos gráficos estatísticos conhecidos hoje foram desenvolvidos nesta época. Na segunda metade do século XIX as teorias desenvolvidas por Carl Friedrich Gauss (Progressões Aritméticas e Geométricas, Método dos Mínimos Quadrados, entre outras) e Pierre Simon Laplace (teoria das Probabilidades, Transformada de Laplace, entre outras) forneceram meios para o entendimento de uma grande quantidade de dados. O início do século XX foi de consolidação e popularização das ideias desenvolvidas no século anterior. Nesta época a visualização gráfica é consagrada para explicar novas ideias e teorias. A grande mudança veio com a chegada do computador, na segunda metade do século 20.

A partir de 1975, com a chegada do computador as inovações e descobertas foram muitas e em diversas áreas. O desenvolvimento de *softwares* altamente interativos e de fácil manipulação vem proporcionando a quebra de novos paradigmas na manipulação e na visualização multidimensional dos dados.

Um exemplo de plataforma de visualização de dados é o *Google Maps*, muito utilizada no cotidiano como ferramenta de localização nas grandes cidades. Além deste uso corriqueiro o *Google Maps* permite, entre outras opções, a sobreposição de dados georeferenciados.

Nos dias atuais, com o aumento da quantidade de informação que pode ser obtida através das diversas redes que nos rodeiam, se torna imprescindível a

visualização e análise dos dados coletados, de forma a realimentar o sistema e permitir uma melhora no desempenho dos mesmos.

Uma destas redes, que se torna cada vez mais popular e presente nas cidades brasileiras é a rede de telefonia celular. Na última década o número de aparelhos celulares no país passou de 23 milhões para 202 milhões. Este dado estatístico nos mostra a importância de se ter um bom gerenciamento e operação das redes celulares pelas empresas concessionárias.

Sou formado em Engenharia Eletrônica e de Telecomunicação pela Puc-MG e atualmente trabalho como Engenheiro de Telecomunicações, na área de projetos de radiofrequência em uma grande operadora de telefonia celular. Facilitar o meu trabalho e de meus colegas é uma das motivações para o desenvolvimento deste trabalho, que tem como objetivo descrever um método de visualização de dados estatísticos, georeferenciados. Os dados trabalhados foram coletados de uma rede de telefonia celular e se referem ao tráfego telefônico cursado na cidade de Varginha - MG nos mais diferentes nós desta rede. Para atingir este objetivo foi utilizado o software gratuito R que permite a visualização dos dados em mapas da plataforma Google Maps.

2 MOTIVAÇÃO

A motivação para este trabalho partiu da necessidade de um novo método para visualização de dados coletados de uma rede de telefonia móvel de uma grande empresa do ramo. Uma rede de telefonia móvel (celular) gera a cada período de coleta (15 minutos) uma infinidade de dados referentes aos mais diferentes tipos de indicadores e elementos envolvidos no processo. Um destes indicadores e sem dúvida um dos mais importantes é o tráfego telefônico cursado por cada uma de suas células, que estão espalhadas pelas cidades, estradas, etc. e que são responsáveis por irradiar o sinal que permite o estabelecimento da conexão do usuário com a rede e por conseqüência com o usuário de destino.

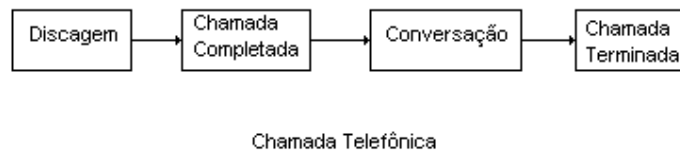


Figura 1 – A chamada telefônica. Fonte: Teleco.com.br

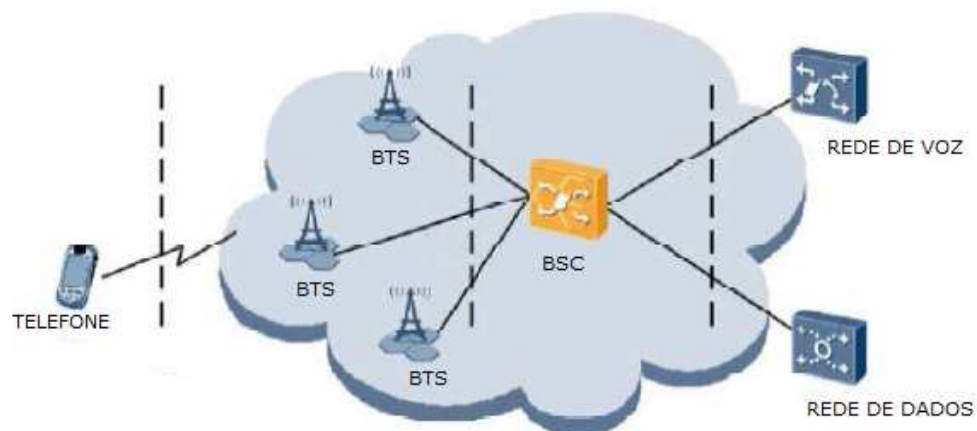


Figura 2 - Topologia de rede celular GSM

A topologia de uma rede celular GSM, mostrada na figura 2, pode ser explicada sucintamente através dos seguintes elementos:

Telefone Celular: dispositivo utilizado pelo usuário para se conectar a rede.

BTS: é a estação propriamente dita, responsável por irradiar o sinal captado pelos aparelhos e estabelecer a conexão do aparelho com a rede.

BSC: é o elemento controlador das diversas estações (BTS) espalhadas pela cidade.

Rede de comutação de chamadas de voz: responsável pela conexão entre as diferentes BSCs dos usuários que utilizam o serviço (assinante A e assinante B).

Rede de comutação de chamadas de dados: responsável pela conexão do usuário com a internet e com outras redes de acordo com a necessidade do cliente.

Cada estação celular (BTS) possui uma estrutura vertical (torre, poste, prédio, etc) onde estão instaladas as antenas responsáveis por irradiar o sinal nas direções desejadas. Normalmente cada estação possui três antenas, que irradiam o sinal em três direções distintas. Cada uma destas antenas estão associadas logicamente no sistema a uma célula da estação, ou seja, podemos dizer que normalmente cada estação possui três células ou setores.

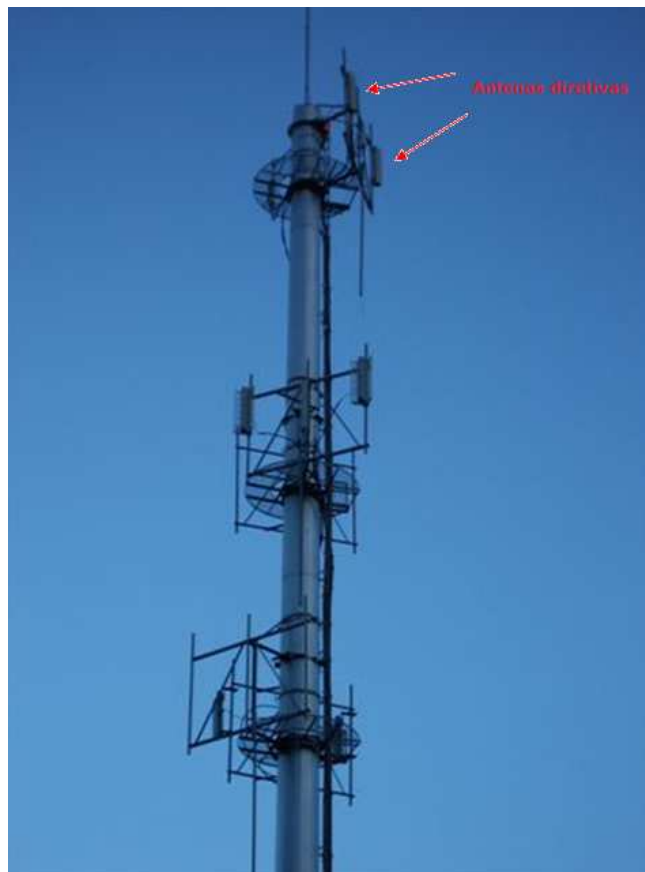


Figura 3 - Torre celular e suas antenas direcionais

A capacidade de uma célula em permitir o acesso e a utilização da rede pelo usuário varia de acordo com vários equipamentos instalados em cada um dos setores. Uma das necessidades de se monitorar o tráfego cursado hora a hora nesta célula justifica-se justamente para que esta célula não chegue no limite e passe a recusar a conexão de novos usuários.

A medida do tráfego telefônico em uma célula é feita através de coletas periódicas de pacotes de informação gerados pela rede, que são enviados para um banco de dados de a cada hora. O valor é reportado na unidade denominada *Erlang*.

3 METODOLOGIA

3.1 A Unidade *Erlang*

Uma rede de telefonia é muito complexa e envolve vários elementos, que tem que ser bem dimensionados e monitorados de forma que se mantenham em operação constantemente, sem interrupções e que esse funcionamento ocorra dentro dos limites do próprio equipamento.

Cada uma das células de uma estação celular é dimensionada para suportar um número máximo de ligações simultâneas. Se chegarmos neste limite passamos a ter congestionamento naquele elemento da rede.

O funcionamento da rede é dinâmico, varia de acordo com os hábitos e necessidades do usuário e devido a essa característica, ao longo do dia poderemos ter grandes variações de tráfego, oscilando entre a ociosidade e o congestionamento.

A complexidade da rede telefônica cria uma dificuldade para os engenheiros responsáveis pela identificação de todos os fatores envolvidos e suas características. Podemos citar, por exemplo, um congestionamento na cidade, gerado por um acidente e que vai acabar produzindo um aumento de tráfego nas estações da região devido a grande concentração de usuários.

A intensidade de tráfego em um sistema telefônico pode ser definida como o somatório dos tempos das chamadas telefônicas (ocupação dos canais telefônicos) em um determinado período de tempo, normalmente de uma hora. A medida de intensidade de tráfego telefônico é o *Erlang* (Erl), considerando o intervalo de uma hora.

Exemplos:

1 canal ocupado durante uma hora representa 1 *Erlang*.

2 canais ocupados durante 30 minutos, somados, também representam 1 *Erlang*

Os dados utilizados neste trabalho são dados de tráfego, em *Erlang*, coletados durante um período de 4 meses, de várias células de uma região.

3.2 Série Temporal Georeferenciada

Neste trabalho, um conjunto de séries temporais está vinculado a coordenadas geográficas, ou seja, podemos afirmar que estas séries são uma função de dois elementos: tempo e espaço.

$$y = f(\textit{tempo}, \textit{espaço})$$

Existem algumas formas de análise de séries temporais. Podemos partir para uma análise dos valores já coletados, estudando algumas características da série como: a periodicidade, os valores máximos e mínimos, a distribuição dos valores, entre outros. Outra forma é tentar descrever matematicamente o comportamento da série, de forma que se possa prever valores futuros ou até mesmo falhas que tenham ocorrido no período de coleta.

Neste trabalho estamos optamos por fazer uma análise descritiva dos dados, propondo um método de visualização georeferenciada das séries, de forma que possamos explorar visualmente a influência das duas variáveis (tempo e espaço) no comportamento das mesmas.

Tratando especificamente dos dados em estudo, podemos verificar que temos algumas valores ausentes nas séries. Esses valores ausentes representam alguns dias em que o sistema de coleta apresentou problemas e, portanto, os dados foram perdidos. Nesses períodos assumimos que os valores para cada uma das séries é zero, ou seja, sem sinal.

4 ESTATÍSTICAS DESCRITIVAS

4.1 Análise da série

Existem várias formas de visualização dos valores de uma série. Abaixo temos a descrição de alguns deles.

4.1.1 *Box Plot*

Num gráfico do tipo Box Plot podemos observar 5 variáveis referente a uma série. São eles:

- **Mediana (Q2):** está representada pela linha mais grossa, dentro do retângulo (caixa ou *box*).

- **Quartil superior (Q3):** é a linha superior do retângulo.

- **Quartil inferior (Q1):** é a linha inferior do retângulo.

- **Cerca superior:** Valor igual a: $Q3 + 1,5(Q3 - Q1)$

- **Cerca inferior:** Valor igual a: $Q1 - 1,5(Q3 - Q1)$

Além disso, temos os pontos com valores discrepantes, que estão fora dos limites dos valores cerca inferior e superior, citados acima. Estes pontos são marcados individualmente no gráfico.

Na figura 4 a representação de uma série em quatro diferentes dias da semana. Os 5 pontos estão mostrados no gráfico.

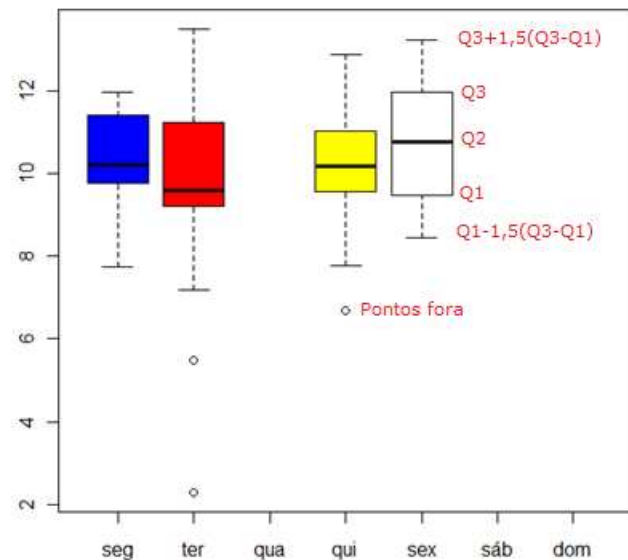


Figura 4 – Exemplo de Box Plot de uma série temporal, extratificado por dia da semana.

4.1.2 Histograma

O histograma é também conhecido como barra de freqüências. É um gráfico de barras verticais em que a altura da barra mostra a freqüência de ocorrência de um intervalo e a base representa o intervalo dos valores medidos.

Na figura 5 temos um exemplo deste tipo de gráfico.

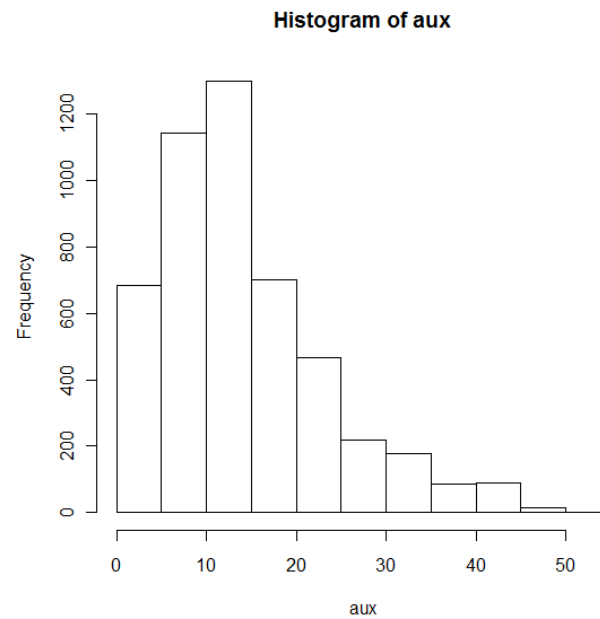


Figura 5 - Exemplo de Histograma

4.1.3 Diagrama Dispersão

O diagrama de dispersão é um gráfico onde pontos no sistema cartesiano XY são usados para representar simultaneamente os valores de duas variáveis, medidas em cada elemento do conjunto de dados.

No exemplo da figura 6 podemos visualizar a quantidade de setores que possuem determinada quantidade de canais de tráfego de voz.

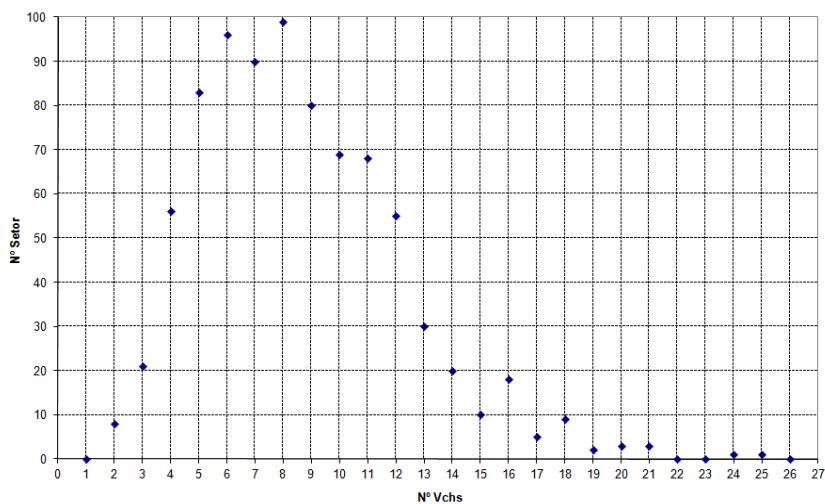


Figura 6 – Diagrama de dispersão

4.1.4 Gráfico de Linha

O gráfico de linha é um gráfico temporal que mostra os valores de determinada variável relacionada com o tempo.

No gráfico abaixo podemos visualizar os dados de tráfego de uma determinada célula do sistema celular, em um período de coleta de dados. A linha vermelha superior indica a capacidade desta célula em Erlang. Veja que a medida que o setor foi chegando perto de sua capacidade máxima ele teve que ser ampliado de forma a evitar congestionamento de rede.

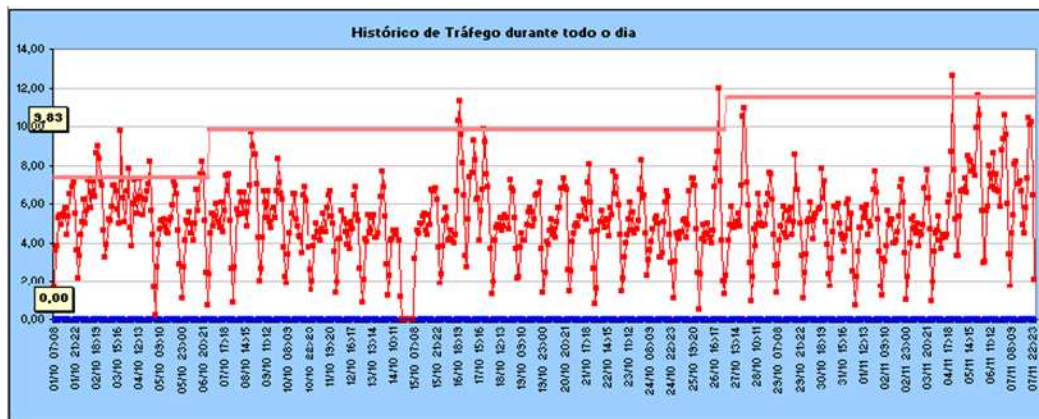


Figura 7 – Gráfico de linha

5 GEOREFERENCIAMENTO

Um grande exemplo da importância do georeferenciamento é o Censo 2010, realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2011) e a forma de coleta dos dados pelos recenseadores. Nas pesquisas anteriores os dados eram todos coletados em formulários, preenchidos a mão, em cada uma das residências do país. Já no último Censo, em 100% das casas a coleta foi feita com pequenos computadores de mão, que já vinculavam a informação digital às coordenadas da residência em questão. Outro exemplo de georeferenciamento, muito utilizado no dia-a-dia, é o site de pesquisa Google Maps (<http://maps.google.com.br>) no qual ao digitar um endereço ele mostra a localização do endereço além de diversos locais de destaque na vizinhança como farmácias, bancos, restaurantes, entre outros.

Conforme exemplificado, com a evolução da tecnologia, fica cada vez mais fácil a manipulação de dados estatísticos e o georeferenciamento dos mesmos, facilitando a análise posterior.

5.1 Sistemas de coordenadas

Existem vários tipos de cálculo das coordenadas geográficas. São eles: as coordenadas astronômicas, as coordenadas geodésicas e as coordenadas naturais. Não é o foco deste trabalho entrar no detalhe de cada um destes modelos, mas apenas definir o formato utilizado neste trabalho, dentro dos diversos sistemas de coordenadas geográficas existentes.

Uma vez que a superfície terrestre (geóide) é irregular, conforme Figura 8, a sua forma pode ser aproximada a partir de uma forma geométrica simplificada: o elipsóide de revolução mostrado na Figura 9.

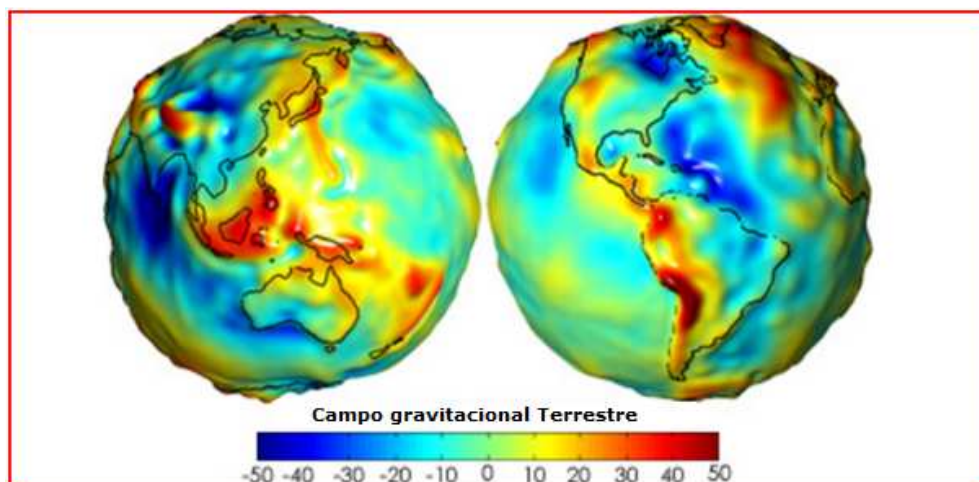


Figura 8 – Geóide: Modelo físico da superfície terrestre (fonte: UFRGS, 2009)

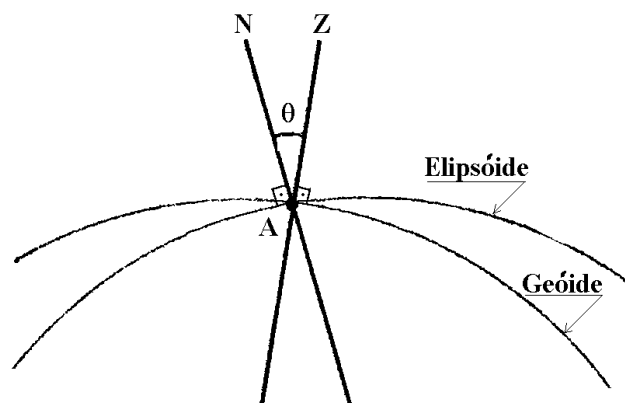


Figura 9 – Elipsóide de revolução

No elipsóide de revolução, o par de coordenadas geográficas é representado pela latitude e longitude de um ponto. A latitude varia de 0° , para pontos sobre a linha do equador, a $+90^\circ$ ou a -90° , estando as latitudes positivas localizadas no hemisfério norte e as latitudes negativas localizadas no hemisfério sul. A latitude geodésica é calculada a partir do ângulo que a reta normal (reta que passa por um ponto do espaço, perpendicular ao elipsóide) ao elipsóide faz com o plano do equador. Essa latitude é diferente da latitude astronômica ou natural que é calculada a partir da reta vertical (reta que passa por um ponto do espaço, perpendicular ao geóide) em relação ao plano do equador (Figura 9).

A longitude de um ponto varia de 0° para pontos sobre o meridiano de Greenwich a $+180^\circ$ ou a -180° , contada positivamente ao leste do meridiano. A

longitude geodésica é calculada pelo ângulo diedro formado pelo plano do meridiano geodésico e o plano do meridiano geodésico local.

As coordenadas de localização de cada uma das estações celular mostradas neste trabalho foram calculadas no Sistema Geodésico WGS 84 (UFSC, 2011), cujas características do elipsóide estão mostradas na tabela abaixo:

Tabela 1 – Elipsóide WGS 1984

SEMI-EIXO MAIOR a	6378137.0 metros
EXCENTRICIDADE	0.00669437999013
SEMI-EIXO MENOR b	6356752.3142 metros

Além do modelo de coordenadas temos o formato de representação das coordenadas geográficas. Existem três modelos, mais utilizados:

- **GMS (Graus - Minutos – Segundos):** Neste sistema, cada grau é dividido em 60 minutos, que por sua vez se subdividem, em 60 segundos. A precisão decimal é definida na casa dos segundos.
Ex: 19°33' 1.64"S // 39°04' 44.06"W
- **Graus - Minutos Decimais:** Cada grau é dividido em 60 minutos. A precisão decimal é dada na casa dos minutos.
Ex: 15°32.192'S // 55°7.491'W.
- **Graus Decimais:** A precisão é definida na casa dos graus. Nesta representação existem valores positivos e negativos. Os valores positivos são para o Norte (latitude) e o Leste (longitude) e não recebem um símbolo específico. Os valores negativos são para o Sul (latitude) e o Oeste (longitude).
Ex: lat -22.906014°lon -47.060571°

Os dados utilizados neste trabalho, foram coletados no formato GMS e convertidos em Graus Decimais no programa R.

5.2 Ferramenta proposta: Programa R

O programa R é um *software* livre que possui uma linguagem programação de mesmo nome (R). É bastante utilizado para análise estatística e a produção dos mais variados tipos de gráficos. Possui uma grande vantagem que é a disponibilização de pacotes específicos para as diversas funcionalidades.

Um desses pacotes é o RGoogleMaps que permite a visualização de pontos na plataforma Google Maps. Este pacote foi utilizado neste trabalho.

5.2.1 Representação da estação (ponto)

A forma geométrica proposta para representação da localização da estação celular é o ponto na coordenada referente a cada uma delas, no mapa georeferenciado do Google Maps. Foi definida uma janela de visualização em que todos os pontos possam ser visualizados. Em seguida, escolheu-se a camada de visualização do mapa, a cor e o tamanho dos pontos a serem plotados.

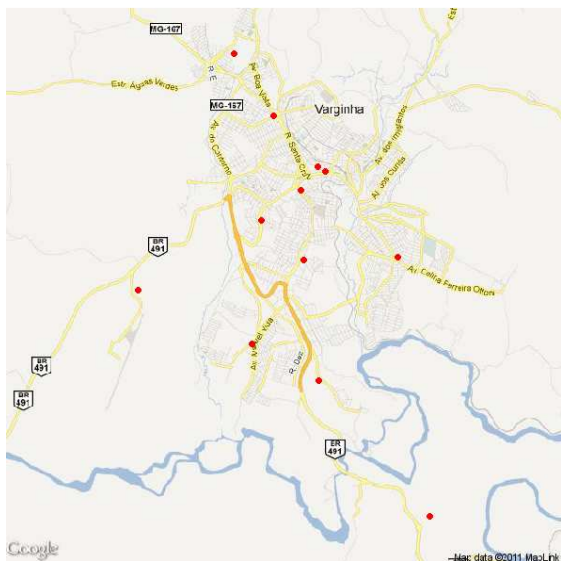


Figura 10(a) - Opção de Visualização: Road Map, ponto vermelho

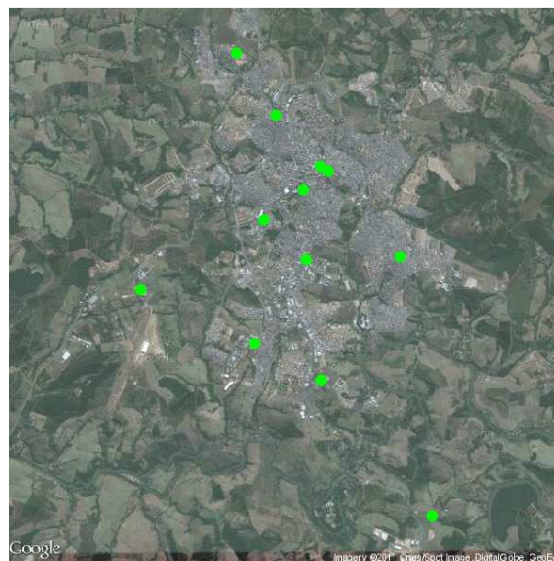


Figura 10(b) - Opção de Visualização: Satellite, ponto variando a cor e o tamanho

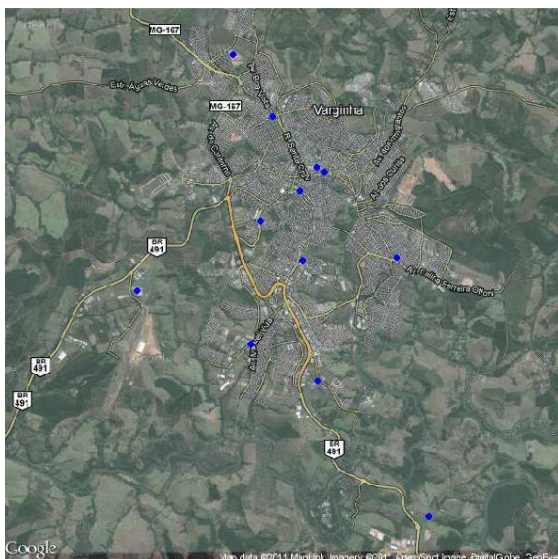


Figura 10(c) - Opção de Visualização:
hybrid = satellite + roadmap, ponto
variando a cor e o tamanho

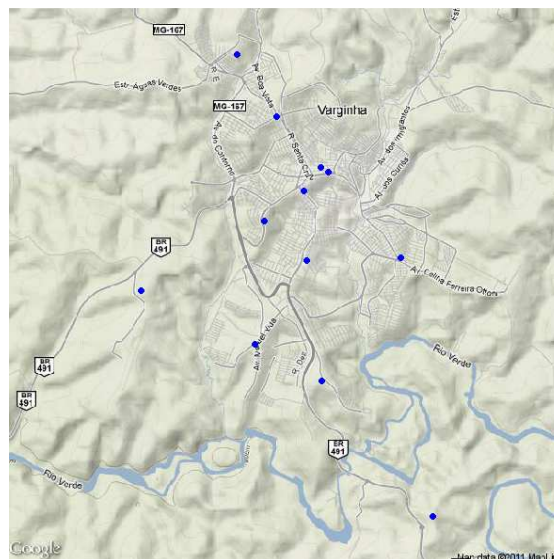


Figura 10 (d) - Opção de Visualização:
Terrain

Figura 10 – Diferentes opções de visualização do mapa, utilizando RGoogleMaps

5.2.2 Representação da Célula (polígono)

Após plotarmos os pontos referentes a cada uma das estações, o passo seguinte foi plotar a visualização de cada um dos setores presentes naquela estação, apontando a direção de cobertura de cada um deles.

Para cada setor de uma estação temos a informação da direção referente àquele setor e a abertura da antena.

A forma mais precisa para desenhar este polígono foi criando um círculo centrado nas coordenadas da estação e plotando no mapa apenas o setor deste círculo referente a direção da antena, com uma abertura de 64 graus.

5.2.2.1 Dinâmica da célula (cor e tamanho)

Uma vez que as estações e cada um dos setores já estão representados no mapa, o passo seguinte foi variar o tamanho e a cor da célula de acordo com a informação de tráfego referente a cada um dos dias coletados.

As janelas de variações foram baseadas nos quartis observados no total de dados coletados. A ordem foi:

- Dados de tráfego concentrados no intervalo que engloba até 20% das amostras;
- Dados de tráfego concentrados no intervalo que engloba entre 20% e 40% das amostras;
- Dados de tráfego concentrados no intervalo que engloba entre 40% e 60% das amostras;
- Dados de tráfego concentrados no intervalo que engloba entre 60% e 80% das amostras;
- Dados de tráfego concentrados no intervalo que engloba mais de 80% das amostras.

As cores utilizadas foram, respectivamente, azul, verde, amarelo, laranja e vermelho e os tamanhos foram sendo variados do menor para o maior, seguindo a própria lógica dos dados.

6 RESULTADOS OBTIDOS

6.1 Localização das estações e células

Na figura 11, mostrada abaixo, temos a representação de cada um dos setores, referentes a cada uma das estações na direção correta e com abertura de 64 graus da antena. Existem setores com sobreposição de cobertura, que estão representado em um vermelho mais escuro.

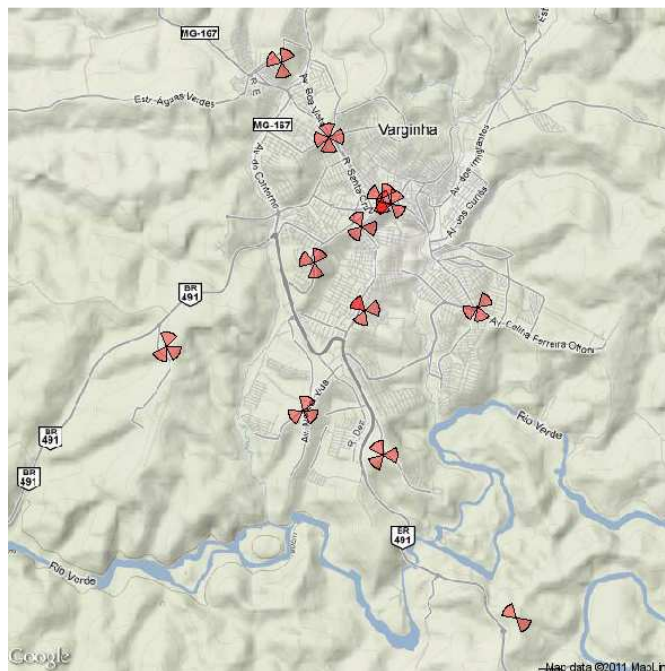


Figura 11 – Células e os respectivos setores, mapa estilo *Terrain*

6.2 Histograma da série

Na figura 12, temos o histograma dos dados de tráfego coletados. Analisando o gráfico observa-se que a maioria dos valores das séries temporais estão concentrados entre 5 e 15 Erlangs.

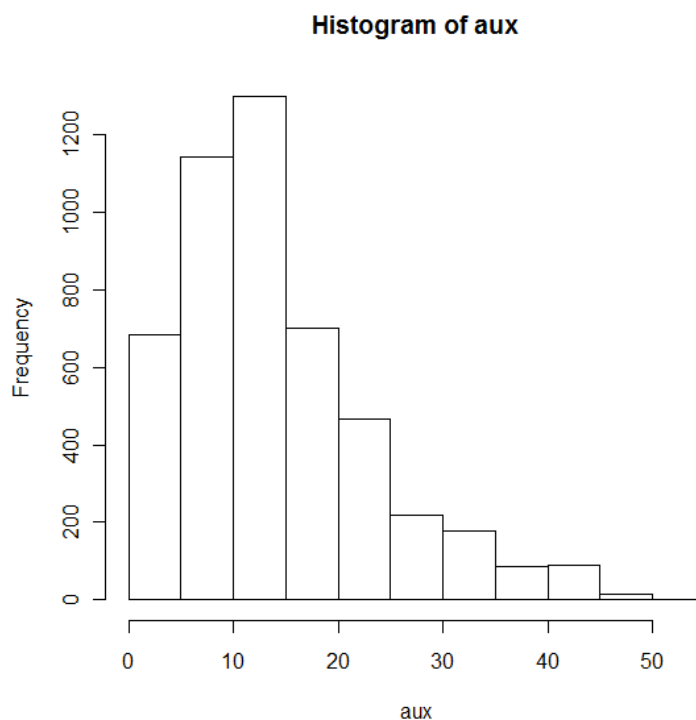


Figura 12 – Histograma dos dados trabalhados

6.3 Gráfico de linha

Na figura abaixo temos o gráfico de linha de todas as séries envolvidas neste trabalho, uma série por setor. Pode-se observar que em todos os casos o tráfego durante os dias da semana é bem maior que o tráfego durante os finais de semana (sábado e domingo). Esse fato, já explicado anteriormente, é influenciado pelo hábito das pessoas na utilização do telefone celular.

No início do período de coleta tivemos algumas falhas de medições, como mostrado nos valores nulos das séries.

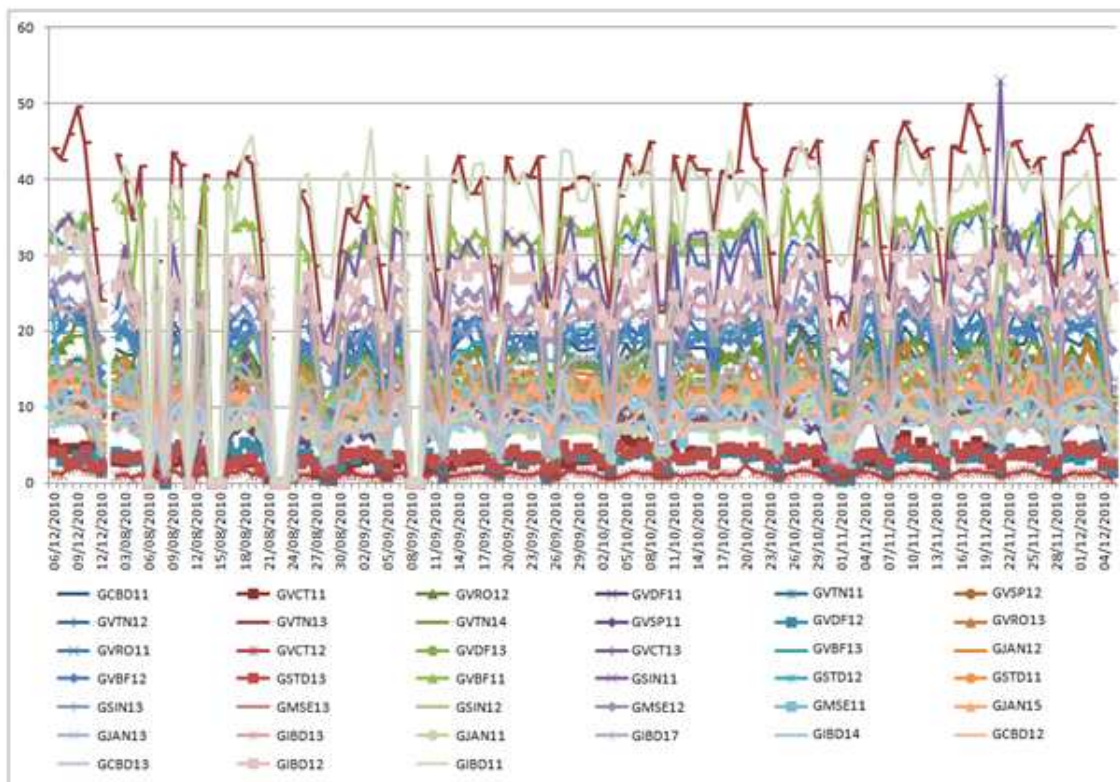


Figura 13 – Gráfico de linha das séries envolvidas

6.4 Comportamento das Séries Temporais por Dia da Semana

Na figura 14, temos o gráfico de Box Plot, para a estação GVTN13. Pode-se observar pelo gráfico que durante a semana o tráfego varia entre 38 e 45 Erl aproximadamente e que nos fins de semana ele vai bastante, chegando a aproximadamente 22 Erl em média no domingo.

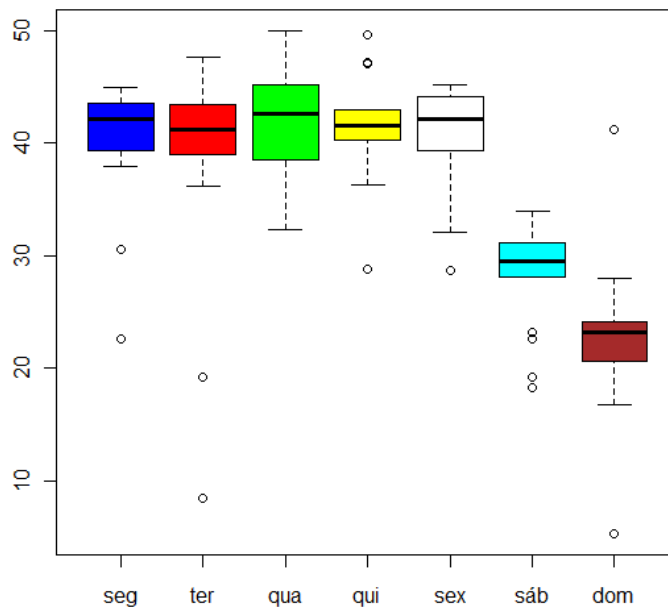


Figura 14 – Gráfico Box Plot setor GVTN13

Este comportamento também é verificado nas demais estações.

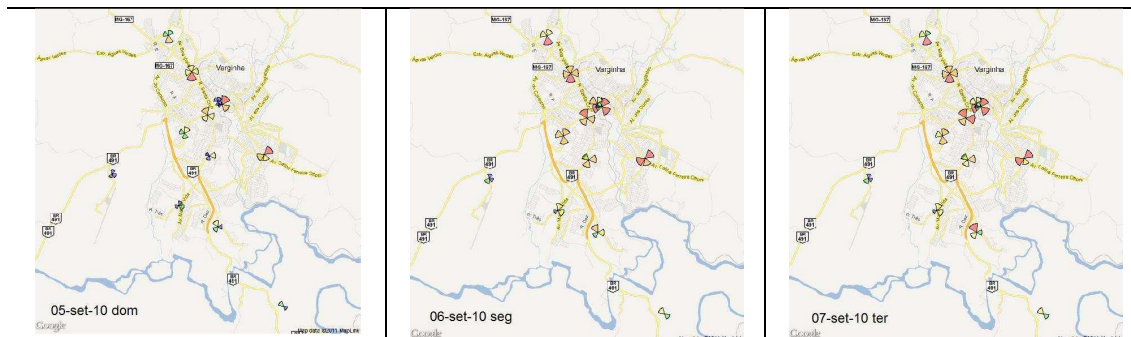
6.5 Análise quadro a quadro (um dia da semana)

Abaixo temos uma animação, quadro a quadro, com a seqüência de quintas feiras pertencentes ao período de análise. A legenda é a seguinte:

- Dados de tráfego concentrados no intervalo que engloba até 20% das amostras **Cor azul, raio do setor: 0,140 Km;**
- Dados de tráfego concentrados no intervalo que engloba entre 20% e 40% das amostras **Cor verde, raio do setor: 0,180 Km;**
- Dados de tráfego concentrados no intervalo que engloba entre 40% e 60% das amostras **Cor amarelo, raio do setor: 0,220 Km;**
- Dados de tráfego concentrados no intervalo que engloba entre 60% e 80% das amostras: **Cor laranja, raio do setor: 0,260 Km;**
- Dados de tráfego concentrados no intervalo que engloba mais de 80% das amostras: **Cor vermelha, raio do setor: 0,3 Km**

Tabela 2 – Sequência de 15 dias





Considerando a seqüência diária de figuras, mostradas na tabela acima, podemos observar o comportamento da população durante os dias da semana e em determinadas regiões da cidade. Existe uma concentração maior de tráfego nas células que cobrem o centro da cidade. Já as células que cobrem a periferia têm menor tráfego. Como em qualquer cidade, o tráfego se concentra nas regiões mais comerciais e com maior fluxo de pessoas.

Tabela 3 – Sequência de quintas feiras





Na tabela 3 temos uma seqüência de imagens de todas as quintas feiras do período de dados trabalhados. Conseguimos perceber claramente a variação de tráfego nos diferentes setores da cidade. Alguns dos fatores avaliados podem ser alterados, por exemplo: alterar o range de variação da cor e tamanho do setor na figura, alterar o período e o método de coleta de dados, entre outros. Uma das conclusões práticas que podemos tirar destas figuras para utilização no dia-a-dia é observar se o tráfego está aumentando em determinada célula com o passar do tempo, de forma a tomar providências necessárias de ampliação de capacidade de rede, evitando assim congestionamentos. Lembrando que no caso desta análise foi escolhido o dia da semana quinta feira porque historicamente é neste dia que normalmente ocorre o maior tráfego da semana. Podemos destacar que todos os indicadores de rede avaliados pela ANATEL (Agência Nacional de Telecomunicações) são coletados às quintas feiras.

7 CONCLUSÃO

O trabalho aqui apresentado buscou descrever um método de análise de dados estatísticos georeferenciados, mostrando sua aplicação no dia-a-dia da engenharia. Na operação diária de uma rede celular, cada vez mais dinâmica e popularizada, em que os usuários querem qualidade de serviço e cobertura, a análise da enorme quantidade de indicadores se torna cada vez mais importante. Em muitas situações a análise dos dados estatísticos norteia a tomada de decisões e o investimento de milhões de reais. Aqui foi descrita uma nova metodologia para visualização dos dados, utilizando uma ferramenta grátis, disponível na internet, acessível para todos.

Podemos destacar também que para análise neste trabalho escolhemos os dados de tráfego (em Erlang) dos setores, mas existem vários outros indicadores de rede em que podemos aplicar este tipo de análise de forma a verificar o desempenho e a tendência de uma rede celular. Entre eles podemos citar: a taxa de queda de chamadas, a taxa de acessibilidade na rede (Sucesso/tentativa de acesso), a utilização de mensagens de texto, o tráfego de dados, a taxa download/upload por setor, entre outros.

8 BIBLIOGRAFIA

TUDE, Eduardo. Tráfego Telefônico (Erlang), 2003. Disponível em: <<http://www.teleco.com.br/tutoriais/tutorialerlang/default.asp>>. Acesso em: 28 jun. 2011.

ERLANG.com.br. Teorias para Dimensionamento, 2001. Disponível em: <<http://www.erlang.com.br/brtheory.asp>>. Acesso em: 28 jun. 2011.

SILVA, Marcos. Boxplot, 2006. Disponível em: <<https://sites.google.com/site/marcosfs2006/boxplot>>. Acesso em: 15 ago. 2011.

ESTATCAMP. Boxplot, 2011. Disponível em: <<http://portalaction.com.br/content/31-boxplot>>. Acesso em: 15 ago. 2011

SHIMAKURA, Sílvia Emiko. O Diagrama de Dispersão, 2002. Disponível em: <<http://www.leg.ufpr.br/~silvia/CE055/node15.html>>. Acesso em: 15 ago. 2011.

CORDINI, Jucilei. Sistema de coordenadas geográficas, 2011. Disponível em: <<http://www.topografia.ufsc.br/cap3-2.html>>. Acesso em: 25 ago. 2011.

CHAFFE, Laureano IBRAHIM. O geóide, 2009. Disponível em: <<http://www.ufrgs.br/museudetopografia/Artigos/Ge%C3%B3ide.pdf>>. Acesso em: 25 ago. 2011.

IBGE, Instituto Brasileiro de Geografia e Estatística. Censo 2010. Disponível em: <<http://www.censo2010.ibge.gov.br/>>. Acesso em: 01 set. 2011.

R, The R Project for Statistical Computing. Disponível em: <<http://www.r-project.org/>>. Acesso em: 15 jun. 2011.

ANEXO 1 – Código R

```

## Código R p/ Georeferenciamento
## Data 19/06/2011
## Marcelo Azevedo/Samuel Contin

rm(list=ls(all=TRUE))
require(PBSmapping)                ## install.packages('PBSmapping', dependencies=TRUE)
require(RgoogleMaps)              ## install.packages('RgoogleMaps', dependencies=TRUE)
## fix(mypolygon)                  ## PARA CORRIGIR BUG NO RGOOGLEMAPS

source("C:\\Users\\Samuel Contin\\_Particular\\Monografia
Pós\\2011_Samuel_Matematica_20110627\\circle_builder.R")
setwd("C:\\Users\\Samuel Contin\\_Particular\\Monografia
Pós\\2011_Samuel_Matematica_20110627")

## Faz a leitura das series de dados
series <- read.table("series_completas.txt",header=TRUE)
## colnames(series)
aux <- series[,3:41]
aux <- (aux[!is.na(aux)])
hist(aux)
range <- quantile(aux,prob=c(0.20,0.40,0.60,0.80))

geo <- read.table("geo.txt",header=TRUE)
geo$lat <- - (geo$LAT_G + geo$LAT_M/60 + geo$LAT_S/3600)
geo$lon <- - (geo$LONG_G + geo$LONG_M/60 + geo$LONG_S/3600)

#####
## PLOTA OS CIRCULOS NO MAPA
## Cluster Primario
zoom <- 13
maptype <- 'roadmap'              ## 'roadmap', 'satellite', 'hybrid', 'terrain'
##maptype <- 'terrain'

n_pix <- 640
center <- c(lat=-21.5676985754986, lon=-45.4408830484330)
destfile <- 'TemporaryMap.jpg'

V_Map <- GetMap.bbox(lonR=c(min(geo$lon),max(geo$lon)), latR=c(min(geo$lat),max(geo$lat)),
  size = c(n_pix, n_pix), zoom = zoom, format = 'jpg',
  maptype = maptype, destfile = destfile)

##for(t in 1:dim(series)[1]){
for(t in 130:130){

  windows()
  PlotOnStaticMap(V_Map)
  ## Plotar os pontos
  PlotOnStaticMap(V_Map,geo$lat,geo$lon,add=TRUE,destfile=destfile,cex=0.5,pch=20,col='blue')
#####
## PLOTA AS ANTENAS COM AS INTENSIDADES, CHAMADA ERLANG - ANALISE EXPLORATORIA
#####

  const <- 180/pi
  R <- 6378.7 ## (km)
  raio <- 0.250 ## (km)
  abertura <- 32 ## (graus)

  for(cont in 3:dim(series)[2]){

    aux <- series[t,cont]

    if(is.na(aux)){ break }

    if(aux > range[4]){ cor.poly <- '#FF000066'; raio <- 0.300 } ## Maior vermelho
    if(aux < range[4]){ cor.poly <- '#FFA50066'; raio <- 0.260 } ## Laranja
    if(aux < range[3]){ cor.poly <- '#FFFF0066'; raio <- 0.220 } ## Amarelo
    if(aux < range[2]){ cor.poly <- '#00FF0066'; raio <- 0.180 } ## Verde

```

```

    if(aux < range[1]){ cor.poly <- '#0000FF66'; raio <- 0.140 }           # Azul

    circulo <- CircleDecimalDegrees(geo$lat[cont-2],geo$lon[cont-
2],radius=raio,np=180,d=6378.7)
    circ.lat <- c(circulo$lat[1:360],circulo$lat[1:360])
    circ.lon <- c(circulo$lon[1:360],circulo$lon[1:360])

    Up.dir <- - geo$DIR[cont-2] + 360 + 90 - abertura
    Lo.dir <- - geo$DIR[cont-2] + 360 + 90 + abertura

    if(Lo.dir < 0){
      Lo.dir <- Lo.dir + 360
      Up.dir <- Up.dir + 360
    }

    ## Somente com direcao de 30o
    x <- c(geo$lon[cont-2],circ.lon[Lo.dir:Up.dir],geo$lon[cont-2])
    y <- c(geo$lat[cont-2],circ.lat[Lo.dir:Up.dir],geo$lat[cont-2])

    ## Gera um polinomio
    tam <- length(x)
    poly <- data.frame(PID=rep(1,tam),POS=c(1:tam),X=x,Y=y)
    poly <- as.PolySet(poly,projection='LL')

    ## Plota o polinomio no MAPA
    PlotPolysOnStaticMap(V_Map, poly, lwd=.25, col=cor.poly, add=T, border='black')
  }

  ## a <- locator(1);
  text(-202.5,-272,paste(series$Data[t],series$Dia_da_Semana[t],sep=' '),cex = 1.5)

}

## #FF000066 (VERMELHO)
## #FFA50066 (LARANJA)
## #FFFF0066 (AMARELO)
## #00FF0066 (VERDE)
## #0000FF66 (AZUL)

```