

UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"

FACULDADE DE CIÊNCIAS - CAMPUS BAURU

DEPARTAMENTO DE COMPUTAÇÃO

BACHARELADO EM SISTEMAS DE INFORMAÇÃO

CAROLINA JUNQUEIRA FERREIRA

JULIANA D'ALESSIO GRANDINI

**GSMART - UMA FERRAMENTA DE GEOMARKETING PARA A
MEDIÇÃO DO TRÁFEGO DE PESSOAS ATRAVÉS DE REDES WI-FI**

BAURU

2017

CAROLINA JUNQUEIRA FERREIRA
JULIANA D’ALESSIO GRANDINI

GSMART - UMA FERRAMENTA DE GEOMARKETING PARA A MEDIÇÃO DO TRÁFEGO DE PESSOAS ATRAVÉS DE REDES WI-FI

Trabalho de Conclusão do Curso de Bacharelado em Sistemas de Informação apresentado ao Departamento de Computação da Faculdade de Ciências da Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP, Câmpus de Bauru.

Orientador: Prof. Dr. Sidnei Bergamaschi

RESUMO

Num mundo hoje plenamente conectado, informar com precisão tornou-se fundamental, especialmente nos negócios: decisões sobre como escolher o melhor ponto para uma loja tornaram-se complexas, dada a quantidade de variáveis a serem consideradas em tal questão. Os conceitos tradicionais em marketing e administração, sozinhos, já se mostram insuficientes para realização desta tarefa. Nesse contexto, processos decisórios empresariais aliaram-se à computação para obter melhores resultados: tecnologias em hardware e software otimizam a coleta de dados e processam a análise dos mesmos com mais rapidez e eficiência. Origina-se, assim, o *geomarketing*, um conjunto de conhecimentos em geolocalização, administração e marketing, atualmente imprescindível no planejamento estratégico. Entretanto, o acesso aos benefícios desta ferramenta ainda restringe-se à grandes e médias organizações, oferecido por consultorias complexas. Pequenos empreendedores ou mesmo estudantes da área de tecnologia e administração raramente tem oportunidade de aplicar ou pesquisar conceitos práticos de *geomarketing* em sua rotina, já que quase não há opções gratuitas disponíveis para tal finalidade. Baseando-se nessa realidade, este trabalho introduz o *geomarketing* na forma de um pequeno sistema para coleta e aferição de tráfego humano em determinadas zonas, utilizando-se de redes Wi-Fi e dispositivos móveis, tecnologias estas baratas e acessíveis, visando promover seu alcance a qualquer interessado no assunto.

Palavras-chave: *Geomarketing*. Tráfego de pessoas. Redes Wi-Fi. Dispositivos Móveis.

ABSTRACT

In a contemporary world fully connected, information accuracy is essential, even more into business: choosing the best store point has become a complex decision, given the amount of variables considered for this matter. Traditional concepts in marketing and administration alone are insufficient to accomplish such a task today. In this context, decision-making has teamed up with computers to achieve better results: hardware and software technologies optimize data collection and process analysis faster and more efficiently. Geomarketing, a body of knowledge comprehending geolocation, administration and marketing, is now an essential into strategic planning. However, benefactions of this tool are still restricted to large and medium-sized organizations, provided by complex consultancies. Small entrepreneurs or even students in technology and administration fields rarely have the opportunity to apply or research practical concepts of geomarketing in their routine, as there are little free sources available to this purpose. Based in such reality, this work introduces geomarketing shapped as a small system to collect and measure human traffic in certain zones, using Wi-Fi networks and mobile devices, cheap and accessible technologies, promoting knowledge to be reached by any interested in the subject.

Keywords: Geomarketing. People Traffic. Wi-Fi Networks. Mobile Devices.

LISTA DE FIGURAS

Figura 1 – Processamento de imagens 2D - Density People Counter	18
Figura 2 – Processamento de imagens 2D e 3D - Axper People Counter	18
Figura 3 – Street Counting - V-Counter	19
Figura 4 – Visitor Counting e Camera Heatmap - V-Counter	19
Figura 5 – Número de dispositivos conectados - Cisco IBSG 2011	20
Figura 6 – Tráfego mundial de dispositivos móveis por tipo	21
Figura 7 – Quantidade de Celulares no Brasil - Março 2017	21
Figura 8 – Detecção de <i>smartphones</i>	22
Figura 9 – Detecção de veículos	23
Figura 10 – Saída após execução de comando Tshark	26
Figura 11 – Código utilizando Express.js	28
Figura 12 – Arquitetura geral do sistema	31
Figura 13 – Diagrama de Fluxo	32
Figura 14 – Trecho de código do sensor	33
Figura 15 – Formatação do dado	34
Figura 16 – <i>Middleware</i> com Express.js	35
Figura 17 – Rotas do Express	36

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

AP	Access Point - Ponto de Acesso
API	Application Programming Interface - Interface de programação de aplicações
CSS	Cascading Style Sheets
CSV	Comma-Separated Values - Valores separados por vírgula
HTML	HyperText Markup Language - Linguagem de Marcação de Hipertexto
HTTP	Hypertext Transfer Protocol - Protocolo de Transferência de Hipertexto
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things - Internet das Coisas
JSON	JavaScript Object Notation - Notação de Objetos JavaScript
MAC	Media Access Control
NIC	Network Interface Card - Placa de Interface de Rede
RFID	Radio-Frequency IDentification - Identificação por radiofrequência
ROO	Return on Objectives - Retorno dos Objetivos (Métrica)
SO	Sistema Operacional
SQL	Structured Query Language - Linguagem de Consulta Estruturada
TI	Tecnologia da Informação
URI	Uniform Resource Identifier - Identificador Uniforme de Recurso
URL	Uniform Resource Locator - Localizador Uniforme de Recurso
Wi-Fi	Marca registrada da Wi-Fi Alliance. Rede local sem fios baseados no padrão IEEE 802.11

SUMÁRIO

1	INTRODUÇÃO	9
1.1	Problema e justificativa	10
1.2	Objetivos	10
1.2.1	Objetivos Gerais	10
1.2.2	Objetivos específicos	10
1.3	Organização do trabalho	10
2	FUNDAMENTAÇÃO TEÓRICA	12
2.1	Geomarketing	12
2.1.1	Relação entre o marketing, geolocalização e estratégias de negócio . . .	12
2.1.2	Como o geomarketing auxilia e agrega valor ao negócio	12
2.1.3	Casos de uso brasileiros	13
2.1.4	O tráfego de pessoas no geomarketing	15
2.2	Ferramentas para contagem de pessoas	15
2.2.1	Métodos de contagem	16
2.2.2	Contagem de pessoas em pesquisas acadêmicas	16
2.2.3	Produtos empresariais para contagem	17
2.2.4	Escolha do método de contagem	18
2.3	Dispositivos móveis	19
2.3.1	Número e uso de dispositivos	19
2.3.2	Dispositivo móvel como objeto identificador	21
2.3.3	Exemplos em detecção de dispositivos através de redes Wi-Fi	22
3	MATERIAIS E MÉTODOS	24
3.1	Tecnologias utilizadas	24
3.1.1	Comunicação Wi-Fi e <i>probe request</i>	24
3.1.2	Modo monitor	25
3.1.3	Tshark	25
3.1.4	Raspberry Pi	26
3.1.5	Node.js	27
3.1.6	MongoDB	27
3.2	Métodos e Etapas	28
4	ARQUITETURA DO PROJETO	31
4.1	Sensor	31
4.2	Node Transformer	33

4.2.1	Formatação do dado	34
4.3	Node Server	35
4.3.1	Deliberando recursos	35
4.3.2	Processando dados	36
4.4	Visualização de Dados	36
5	RESULTADOS E DISCUSSÃO	37
5.1	Testes e validação do projeto	37
5.2	Análise de Riscos	38
6	CONCLUSÃO	40
	REFERÊNCIAS	41

1 INTRODUÇÃO

Desde os primeiros sistemas comerciais estabelecidos até a Revolução Industrial, a maneira de se fazer negócios evoluiu drasticamente - entretanto, nada se compara ao advento da computação nas empresas: informatizar organizações permitiu coletar e armazenar volumes dados gigantescos, gerando uma quantidade de informação humanamente impossível de ser manipulada. Como consequência, foi imprescindível desenvolverem-se ferramentas e métodos de mineração para que, dentre esse emaranhado de dados, fosse possível extrair-se informação relevante.

Hoje, o conhecimento estatístico é um elemento vital na tomada de decisões em negócios: a qualidade e abrangência das informações levantadas pode significar o sucesso ou fracasso na estratégia empresarial. Avaliar a expansão através de filiais, por exemplo, é um passo arriscado - se mal planejado, poderá incorrer em sérios prejuízos. Estudos detalhados acerca de geolocalização devem ser realizados numa região antes da instalação de uma nova loja.

Ressaltamos aqui, segundo Kotler e Keller (2013), a importância da geolocalização: as atividades de marketing foram classificadas em um mix contendo quatro tipos amplos, denominados *Quatro Ps do marketing*: produto, preço, praça e promoção (do inglês, *product, price, place e promotion*). Neste trabalho, a *praça* é de especial importância: engloba fatores relacionados à geolocalização que influenciam direta e indiretamente nas estratégias de marketing, tais como: canais de distribuição, cobertura, sortimentos, locais, estoque e transporte.

Segundo Mangini, Luz e Conejero (2014), a tomada de decisão em termos de localização não pode ser feita de maneira aleatória e subjetiva, mas embasada em um método ou ferramenta que permita determinar o melhor ponto ou o mais adequado, de acordo com premissas objetivas e dentro de um arcabouço lógico, considerando as possíveis variáveis que afetam aspectos relacionados ao usuário, urbanismo e também relacionado à gestão e às políticas públicas.

Diante deste contexto, o *geomarketing* surge como grande tendência na gestão: reunindo conceitos em geografia espacial, estatística, gestão e marketing, tornou-se a mais abrangente ferramenta para visualização e análise do negócio como um todo - permite explorar vantagens locais, além de gerar um panorama completo de todas as camadas de uma instituição, identificando pontos fortes e fracos e auxiliando na definição das melhores estratégias de planejamento e decisão.

Neste trabalho, o foco do *geomarketing* será voltado à exploração do tráfego de pessoas através da criação do GSmart, uma ferramenta de contagem aliada às redes de

Internet sem fio e dispositivos móveis.

1.1 Problema e justificativa

O tráfego de pessoas já é largamente utilizado como técnica de *geomarketing*, porém no mercado predominam os softwares privados. Além disso, a maioria das ferramentas de contagem utiliza processamento de imagens, tecnologia de alto custo e demasiadamente sofisticada para aplicações mais simples.

Visa-se, assim, criar uma solução em software livre, propiciando uma ferramenta gratuita e de fácil utilização voltada a profissionais da área de marketing, computação, administração, empresas e negócios, etc. Para tal, foram escolhidos como foco redes Wi-Fi e dispositivos móveis (celulares, tablets, etc) - isso se justifica por serem tecnologias acessíveis, relativamente baratas e amplamente utilizadas por qualquer pessoa hoje em dia. Além disso, aplicações de *geomarketing* baseadas em redes sem fio são ainda pouco exploradas em pesquisas acadêmicas.

1.2 Objetivos

1.2.1 Objetivos Gerais

Este trabalho tem como objetivo desenvolver um sistema que mede o tráfego de pessoas em determinadas zonas através de rede Wi-Fi e dispositivos móveis.

1.2.2 Objetivos específicos

- Estudar ferramentas de contagem de pessoas em ambientes e definir tecnologias para identificação e fornecimento de dados de usuários;
- Identificar a tecnologia responsável pela contagem de pessoas;
- Definir o modo como o número de indivíduos será agrupado para gerar o tráfego, indicar como os dados capturados serão agrupados e implementar interface para visualização dos dados gerados;

1.3 Organização do trabalho

O presente trabalho divide-se em capítulos, sendo este o primeiro (Introdução). Os próximos seguirão a ordem abaixo:

- **Fundamentação Teórica:** apresentação dos conceitos teóricos envolvidos no trabalho, casos de uso e soluções semelhantes;
- **Materiais e Métodos:** ferramentas escolhidas para o desenvolvimento;
- **Arquitetura do Projeto:** sensor, preparação de dados, processos, servidor;
- **Resultados e Discussão:** testes, análise de riscos;

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Geomarketing

Segundo Aragão (2005), *geomarketing* é o nome dado à área de gerenciamento de informação que incorpora as dimensões espaciais para auxiliar a tomada de decisões dentro de um domínio específico de mercado, o que permite levantar as características de uma determinada região e analisar seu potencial socioeconômico. Pode ser entendido, assim, como uma ferramenta de análise estatística de dados, com intuito de localizar padrões que possam ser utilizados e combinados na elaboração de indicadores, perfis de consumo e estratégias de negócios, de modo a gerar informação relevante na tomada de decisões. Geralmente, o serviço é oferecido por consultorias especializadas - o objetivo da empresa contratante é a melhoria no desempenho de seu negócio.

2.1.1 Relação entre o marketing, geolocalização e estratégias de negócio

Segundo Cliquet (2006), alguns autores definem o *geomarketing* como uma aplicação específica do espaço econômico, codificando técnicas e divisões geográficas associadas a funções estatísticas. Alguns aspectos espaciais provém ainda características das quais podem ser inferidos novos dados, mais complexos e específicos, a partir de análises mais aprofundadas nos dados iniciais. Este trabalho analítico transforma números em valiosas informações estatísticas, auxiliando na tomada de decisão ao permitir que um gestor visualize de forma mais tangível determinadas situações - por exemplo, projetando estatisticamente cenários positivos ou negativos no futuro do negócio.

Ainda para Cliquet (2006), o espaço geográfico raramente é levado em conta em pesquisas de gestão, exceto quando se trata de escolher a melhor localização para um ponto de vendas ou instalação de uma unidade produtora. No entanto, aspectos espaciais de uma organização são muito mais abrangentes e tem impacto vital no desempenho dos negócios como um todo.

2.1.2 Como o geomarketing auxilia e agrega valor ao negócio

O *geomarketing* tem como principal função servir como base de auxílio quando um administrador se depara com questões cruciais para a continuidade dos negócios, sendo de grande importância na definição do público-alvo ou onde estabelecer novas filiais.

Segundo Duarte (2016), a localização geográfica visualiza de maneira diferenciada as seguintes informações:

- Distribuição de unidades e clientes;
- Distâncias percorridas no consumo dos clientes;
- Padrões de consumo;
- Aspectos no entorno das lojas;
- Características sociodemográficas;
- Pontos interessantes;
- Mapear força de vendas;
- Localização de prospectos e distribuidores.

O grande diferencial em se utilizar o *geomarketing* nos negócios é justamente a eficiência na gestão de informação: transformar dados e números em conhecimento tangível desenvolve a inteligência de uma organização. Sob tal contexto, podemos ainda mencionar que, segundo Duarte (2016), o *geomarketing* traz uma enorme valorização nos pontos mencionados abaixo:

- Otimiza a rede de negócios e norteia a busca de novos pontos;
- Conhece cada loja para análise e comparação;
- Estabelece metas;
- Ajusta os produtos de cada unidade;
- Distribui adequadamente a força em vendas;
- Elimina sobreposição de vendedores;
- Mede o potencial do mercado;
- Analisa a distribuição da concorrência.

2.1.3 Casos de uso brasileiros

O termo *geomarketing* ainda não é tão difundido no Brasil, porém cada vez mais se populariza no âmbito dos negócios: inicialmente, estudos com finalidades geo-estatísticas eram realizados de forma amadora - ao longo dos anos, a super valorização da informação e o desenvolvimento tecnológico fizeram com que a prática evoluísse, juntamente com a especialização de profissionais na área.

Atualmente, o *geomarketing* é parte intrínseca na estratégia de planejamento em organizações brasileiras de todos os tamanhos. Grandes grupos como O Boticário usam largamente o marketing geográfico, assim como pequenas e médias empresas já focam seus sistemas na geolocalização.

Podemos citar como exemplo de aplicação de pequeno negócio um restaurante voltado à alimentação saudável na cidade de Natal - o objetivo do estudo geográfico foi verificar a distribuição de clientes e mapear áreas de influência para conhecer melhor a demanda do mercado. De acordo com Seabra (2014), esta investigação permitiu uma compreensão do fenômeno da área de influência e de variáveis que modelam seu comportamento. O estudo baseou-se em informações obtidas através de softwares como Google Maps para o georreferenciamento e análise dos dados - isso só foi possível graças a fácil disponibilidade e barateamento da tecnologia atual: o Google Maps é um exemplo de ferramenta de geolocalização bastante popular e acessível que, há alguns anos, não existia.

O crescimento acelerado de grandes centros urbanos também gerou uma infinidade de aplicações em *geomarketing*, tornando a ferramenta cada vez mais ampla e complexa - no Brasil, não foi diferente. Um caso de uso a ser citado nesse contexto é a utilização do *geomarketing* como ferramenta de análise para criação de novas estações na CPTM (Companhia Paulista de Trens Metropolitanos). Segundo Mangini, Luz e Conejero (2014), o modelo apresentou ser de grande valia por reduzir de forma substancial a subjetividade da escolha do local para uma nova estação e pôde ainda ser utilizado como método para a definição de novas linhas férreas.

Outro caso de uso interessante é o do Shopping Cidade Jardim, em São Paulo, onde em 2016 a empresa Zebra Technologies implantou seu projeto MPact: utilizando redes Wi-Fi/Bluetooth e oferecendo acesso gratuito à Internet, este captura a localização do cliente em três níveis: zona, posição e presença. Isso permite saber onde ele está, quanto tempo fica em cada setor ou quais produtos está adquirindo, propiciando que varejistas, lojistas e operadores entendam melhor o comportamento dos consumidores. É possível ainda identificar corredores mais cheios, lojas que mais vendem e pontos que recebem maior atenção, auxiliando no monitoramento de vendas. Segundo a Zebra Technologies (2016), esta é uma maneira de compreender o que os clientes querem para, assim, ganhá-los e mantê-los.

Diante do exposto nessa seção, podemos perceber a importância do *geomarketing* como referencial na tomada de decisões estratégicas empresariais, tornando-o hoje uma ferramenta indispensável nos negócios.

2.1.4 O tráfego de pessoas no geomarketing

Segundo ShopperTrak (2017), sem dados de tráfego, estamos apenas tentando adivinhar. Saber quantas pessoas entram numa loja é apenas o início. Quando conhecemos de fato o cliente, percebemos o panorama do que realmente está acontecendo nos negócios.

Conforme Seed (2017), através da contagem de pessoas é possível identificar os períodos de alta e baixa no movimento de clientes ao longo dos horários e dias da semana dentro do estabelecimento. Também é possível calcular a taxa de conversão de negócio, indicando a quantidade de clientes que saíram de uma loja sem efetivar compras.

Ainda para Seed (2017), analisar o fluxo de pessoas permite conhecer o real potencial e aproveitamento do negócio em função do tamanho da loja. Conhecendo o volume de clientes que normalmente ocupa um espaço ao mesmo tempo serve de base para dimensionar a equipe de atendimento, disposição e mix de produtos. Pode-se ainda avaliar a entrada de novos pontos estratégicos de atuação ou mesmo incrementar o alcance nos locais já existentes.

Outros exemplos de dados segundo Seed (2017) e ShopperTrak (2017) que podem ser inferidos a partir da contagem:

- Número de Visitantes por hora;
- Horários de pico por baixa;
- Comparação de fluxo por horários ou por períodos;
- Comparação de performance;
- Valor vendas por Número vendas (Ticket médio);
- Vendas por Visitantes (Taxa de Conversão);
- ROO – Taxa de conversão por ações de marketing;
- Determinar custo adicional por cliente e retorno de investimento em marketing;
- Escala de funcionários otimizada;
- Taxa de atração e tráfego em tempo real.

2.2 Ferramentas para contagem de pessoas

Para que o *geomarketing* e o tráfego de pessoas sejam implementados, é necessário medir o número de pessoas através de uma ferramenta de contagem.

As ferramentas de contagens são sistemas eletrônicos que utilizam leitores para contar as pessoas (TRAF-SYS, 2017). O tráfego é gerado por essa contagem durante certos períodos de tempo. Estes dados, quando aliados a outras métricas de negócio, geram muitas informações estratégicas.

2.2.1 Métodos de contagem

Não existe apenas um método para contar o número de pessoas. As principais diferenças entre os contadores estão em: área de cobertura, volume e tecnologia utilizada. Segundo Ipsos Retail Performance (2015), os principais meios de contagem são:

- **Feixes infravermelhos:** são colocados na entrada de lojas emitindo um feixe infravermelho entre os seus extremos, quando alguém interrompe o feixe, uma entrada é contada. A área de cobertura é pequena e o volume de pessoas que ele permite passando pela porta ao mesmo tempo é baixíssima;
- **Câmeras térmicas:** o uso de sensores térmicos e processamento de imagens. Normalmente, são posicionados no teto para que a imagem capture a temperatura das pessoas e compare com a do ambiente. Este sistema permite alto volume de tráfego e instalação em entradas complexas;
- **Vídeo:** Utilização de algoritmos complexos, inteligência artificial e o processamento de imagens (2D e 3D). A área de cobertura pode ser medida de acordo com o uso de câmeras e o volume permitido varia de acordo com os algoritmos;
- **Wi-Fi:** utiliza o receptor Wi-Fi para pegar *frames* únicos de gerenciamento Wi-Fi emitidos por dispositivos dentro do alcance. Ideal para áreas onde o volume de pessoas é esparsa ou incerto.

A escolha de um contador varia de acordo com a complexidade da entrada do lugar, períodos de captura do tráfego de pessoas, volume de pessoas por período, área de cobertura, precisão desejada, preço, entre outros (WADSWORTH, 2013) (AXPER, 2017).

2.2.2 Contagem de pessoas em pesquisas acadêmicas

Citaremos aqui alguns exemplos de pesquisas acadêmicas focadas em contagem de pessoas. Cada tópico a seguir apresenta um projeto - todos os pesquisadores empregam como ferramentas câmeras de gravação e processamento de imagens. As pesquisas diferenciam-se pelas técnicas de computação utilizadas:

- **Robusto e leve:** com o objetivo de fornecer segurança para ambientes internos o trabalho de Kim et al. (2002) preza por um sistema que seja robusto suficiente

para garantir as metas, mas não seja tão pesado do ponto de vista de algoritmos e demanda de hardware. O sistema reconhece o movimento de pessoas ao longo de várias direções através de uma única câmera e um processador Pentium IV, assim ele estima e rastreia uma "caixa" ao redor de cada indivíduo para identificá-lo na imagem;

- **Melhora no processamento de imagens e ruídos:** as pesquisas de Luo et al. (2016) e Hou e Pang (2011) consideram a queda de desempenho de sistemas de contagem em ambientes com multidões, oclusões (sombreamento/luminosidade em cada quadro do vídeo) e informações de fundo complexas. O primeiro artigo propõe uma abordagem de cenas *indoor* que leva em conta multidões estacionárias (paradas) ou em movimento. O sistema detecta a multidão e separa os ruídos. Depois, estima-se o número de pessoas através de "ombro-cabeça". Por fim, para reduzir as oclusões, há um filtro que separa quadro por quadro do vídeo e faz um tratamento. Já o segundo, foca em subtrair o fundo, estima o número de pessoas e utiliza técnicas para identificar as pessoas em imagens de baixa resolução;
- **Múltiplos recursos:** os artigos de Venkatesh Ankan Bansal (2015) e Ma, Zeng e Ling (2012) consideram múltiplos recursos para contar pessoas em ambientes densos. O primeiro utiliza, principalmente, técnicas matemáticas e técnicas de filtros e imagens para estimar. Já o segundo, utiliza múltiplas câmeras e vários níveis de textura para lidar com aparência humana e posições.

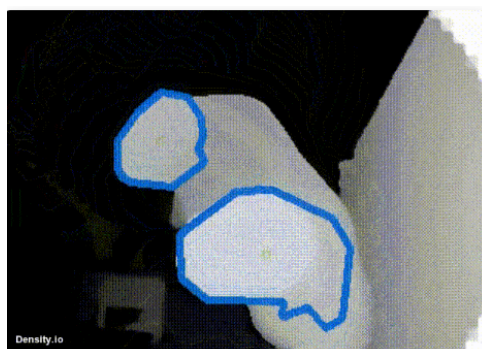
As principais características em sistemas de contagem que os artigos levantados focaram são: movimentação das pessoas, ambientes de multidão e processamento em tempo real.

2.2.3 Produtos empresariais para contagem

Esta seção descreve exemplos de contadores desenvolvidos por empresas que são mais comumente encontrados no mercado:

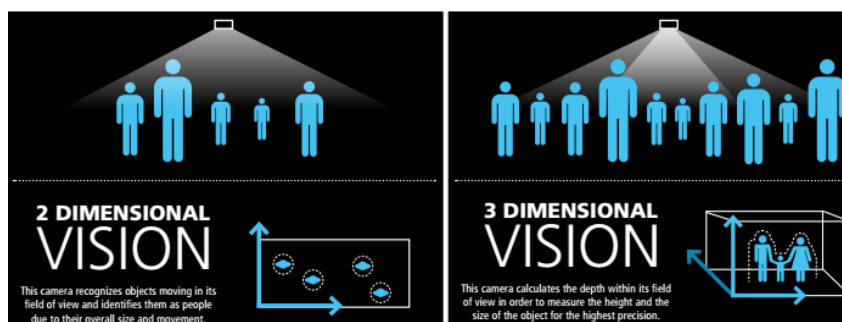
- A empresa Density oferece a contagem a partir de um dispositivo localizado no topo da entrada que processa imagens 2D (DENSITY, 2017), conforme pode ser verificado na Figura 1.
- A Axper, além de oferecer o processamento de imagens 2D, como a Density, disponibiliza também um dispositivo que processa imagens em 3D, realizando a cobertura de todo o ambiente (AXPER, 2017), como pode ser visto na Figura 2.
- A V-Count (2017) oferece soluções de contagem a partir de imagens térmicas e sinais Wi-Fi. Na Figura 3, a contagem ocorre por um aparelho que, fixado na entrada da loja, capta sinais emitidos pelos dispositivos móveis pertencentes às pessoas que

Figura 1 – Processamento de imagens 2D - Density People Counter



Fonte: Density (2017).

Figura 2 – Processamento de imagens 2D e 3D - Axper People Counter



Fonte: Axper (2017).

transitam dentro da área coberta. Já a Figura 4 mostra as soluções que processam imagens e a temperatura para identificar os clientes e seus hábitos.

2.2.4 Escolha do método de contagem

Foi observado que as principais pesquisas de técnicas para a contagem de pessoas abordam o processamento de imagens e vídeo. Já no âmbito empresarial, há diversificadas soluções partindo desde o uso dessas imagens até o uso de emissão de sinais Wi-Fi. As soluções em TI que empregam as redes sem fio para identificar pessoas são variadas. Por exemplo, uma área amplamente explorada em pesquisas é a localização de pessoas em ambientes fechados (*indoor location*) que utiliza dispositivos móveis e emissão de sinais *wireless* (FERREIRA, 2016) (PUHL, 2016) (FIGUERA et al., 2011). No entanto, o uso da técnica de contagem por Wi-Fi como ferramenta do *geomarketing* não é amplamente desenvolvida na comunidade aberta, permanecendo restrita a empresas de consultoria e serviços de TI, como citadas na subseção 2.2.3.

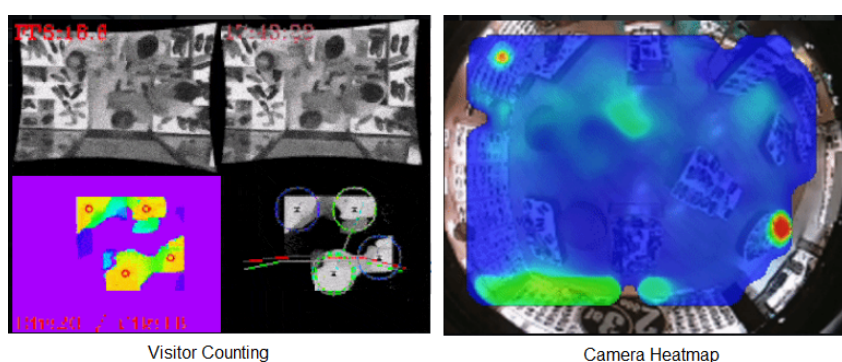
Devido à grande área de cobertura na tecnologia Wi-Fi e à baixa exploração de seu

Figura 3 – Street Counting - V-Counter



Fonte: V-Count (2017).

Figura 4 – Visitor Counting e Camera Heatmap - V-Counter



Fonte: V-Count (2017).

uso como ferramenta de *geomarketing* em comunidade aberta, o presente trabalho priorizou esse tipo de comunicação para a medição do tráfego de pessoas. A escolha considerou ainda o fato de que a identificação de indivíduos em multidões foi apontada como tendência de pesquisa.

2.3 Dispositivos móveis

2.3.1 Número e uso de dispositivos

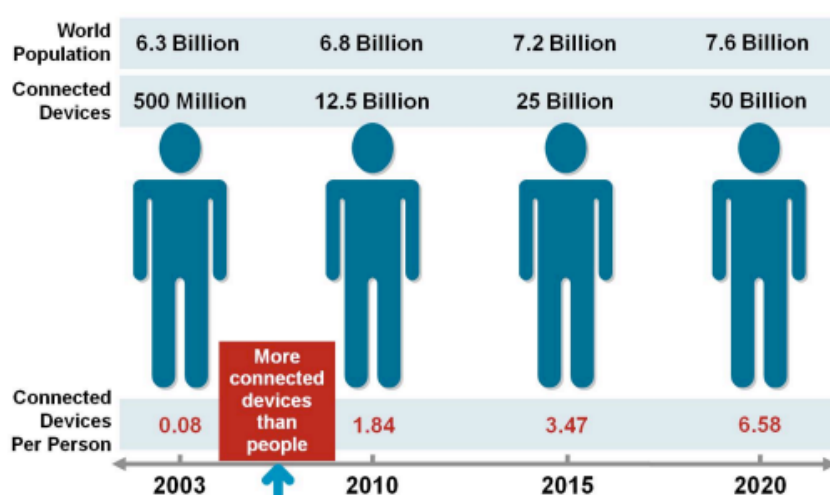
Hoje em dia, é muito difícil encontrar nas ruas uma pessoa que não esteja portando qualquer tipo de dispositivo móvel com acesso à rede - essa proliferação, segundo diversas estimativas abaixo apresentadas, só tende a aumentar - isso torna o uso de redes de

internet e aparelhos móveis uma fonte para exploração em *geomarketing* cada vez mais importante:

- A Gartner (2014) calculou que, até o final de 2016, possuíamos 6,4 bilhões de dispositivos conectados à internet no mundo inteiro, prevendo que em 2020 serão 20.8 milhões de aparelhos em atividade.
- A Cisco Blogs (2013) estimou que, até 2013, havia 80 dispositivos se conectando na Internet por segundo - em 2020, estima-se que atingiremos 250 aparelhos acessando a rede por segundo e existirão, em média, 50 bilhões de dispositivos ativos.

Abaixo, a Figura 5 apresenta estimativas de que, até 2020, teremos quase 7 aparelhos conectados por pessoa:

Figura 5 – Número de dispositivos conectados - Cisco IBSG 2011



Fonte: Evans (2011).

Apesar da expressiva diferença entre as estimativas anteriores, é bastante perceptível o crescimento no uso da rede em dispositivos móveis. A pesquisa da Cisco (2017) mostra o aumento do tráfego de dados nos tipos aparelhos mais utilizados, como pode ser visto na Figura 6. De 2016 até 2021, a elevação em consumo de exabytes por mês se destaca entre os *smartphones*.

No Brasil, segundo a Teleco (2017), a pesquisa da IDC no último trimestre de 2016 calculou que foram vendidos 48,4 milhões de dispositivos no país, sendo 4,9 milhões de celulares tradicionais e 43,5 milhões de *smartphones* - apesar da queda nos últimos anos, a proporção em vendas ainda é maior para aparelhos do tipo *smartphone*.

Já o total de celulares, em março de 2017, era de aproximadamente 242 milhões, como demonstra a Figura 7.

Figura 6 – Tráfego mundial de dispositivos móveis por tipo



Fonte: Cisco (2017).

Figura 7 – Quantidade de Celulares no Brasil - Março 2017

Milhares	Mar/16	Dez/16	Fev/17	Mar/17
Celulares	257.811	244.067	242.919	242.790
Pré-pago	71,41%	67,48%	67,02%	66,85%
Densidade*	125,42	118,04	117,34	117,20
Crescimento Mês	-252	-4.381	-500	-129
	-0,10%	-1,76%	-0,21%	-0,05%
Crescimento Ano	-4	-13.748	-1.148	-1.276
	0,001%	-5,3%	-0,5%	-0,5%
Crescimento em 1 ano	-25.609	-13.748	-15.144	-15.020
	-9,0%	-5,3%	-5,9%	-5,8%

Nota: celulares ativos na operadora. Densidade calculada com a projeção de população do IBGE (Rev. 2013) para o mês respectivo.

Fonte: Teleco (2017).

2.3.2 Dispositivo móvel como objeto identificador

Assim como os RFIDs passivos utilizados em uma cadeia de suprimentos para identificar produtos que vão num caminhão, ou como os códigos de barras que fornecem ao operador de caixa informações sobre um produto, os dispositivos móveis, como os *smartphones*, serão utilizados como objetos identificadores neste trabalho. Tendo em vista que a proporção de *smartphones* é maior que a de celulares tradicionais no Brasil e, em todo o mundo, considera-se esse equipamento de uso comum. Além disso, esses dispositivos móveis fornecem informação pública passiva através da rede Wi-Fi e independem da fabricante para serem detectados numa rede - serão utilizados como meio de identificação das pessoas.

2.3.3 Exemplos em detecção de dispositivos através de redes Wi-Fi

Este subcapítulo apresenta alguns trabalhos que empregam a detecção de dispositivos móveis por sinais de Wi-Fi, utilizando-os como ferramenta de *geomarketing*.

- Meshlium Xtreme: trata-se de um produto da empresa Libelium que detecta dispositivos móveis e veículos através de sinais Wi-Fi e Bluetooth, visando otimizar a inteligência nos negócios. O sistema conta pessoas e automóveis, gerando informações (Libelium Comunicaciones Distribuidas S.L, 2017) como: quantidade de pessoas passando numa mesma rua diariamente, o tempo médio em que permanecem nela, diferenciação entre indivíduos visitantes ou residentes (moradores da região), rotas caminhadas entre lojas, número de veículos em tempo real, o tempo médio que um veículo fica parado, a sua velocidade média e, quando um congestionamento é detectado, o tempo gasto em rotas alternativas.

Os dispositivos móveis não precisam estar conectados a qualquer AP para serem detectados, e podem ser manufacturados por qualquer fabricante. Os veículos são identificados dentro de velocidades até 100 km/h. O objetivo do produto é medir a quantidade de pessoas e carros em determinado ponto e hora específica, possibilitando criar estratégias de negócios com respeito ao tráfego de indivíduos e carros na área designada. As figuras Figura 8 e Figura 9 demonstram o funcionamento do produto.

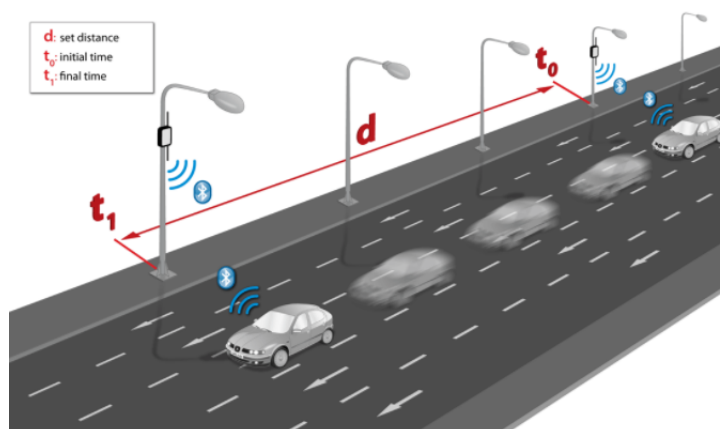
Figura 8 – Detecção de *smartphones*



Fonte: Libelium Comunicaciones Distribuidas S.L (2017).

- How many people are around “How many people are around” (em português: quantas pessoas estão ao redor) é um projeto encontrado no Github do usuário Schollz (2017), que utiliza um *cluster* de Raspberry Pi’s para calcular o número de pessoas próximas e/ou dentro de casa. Para tanto, ele utiliza o protocolo de análise Tshark para detectar

Figura 9 – Detecção de veículos



Fonte: Libelium Comunicaciones Distribuidas S.L (2017).

Wi-Fi *probe requests* de *smartphones* e a linguagem Python. Além disso, os dados capturados podem ser visualizados em forma de gráfico.

3 MATERIAIS E MÉTODOS

Este capítulo descreve as tecnologias e etapas para o desenvolvimento do sistema proposto neste trabalho.

3.1 Tecnologias utilizadas

3.1.1 Comunicação Wi-Fi e *probe request*

Segundo Teleco (2008), uma Wireless LAN (WLAN) é uma rede local sem fio padronizada pelo IEEE 802.11. É conhecida também pelo nome de Wi-Fi, abreviatura de *wireless fidelity* (fidelidade sem fios) e marca registrada pertencente à Wireless Ethernet Compatibility Alliance (WECA).

De acordo com SIMÕES (2015), a tecnologia Wi-Fi é muito utilizada em sistemas de posicionamento indoor, pois não existe a necessidade de criar uma infraestrutura de comunicação: em praticamente todos os espaços fechados com afluência de pessoas, existe já uma criada. Vale ressaltar, porém, que redes Wi-Fi estão sujeitas a pontos cegos (desvios de sinal em áreas não atingidas pela rede) - isso pode ser causado por interferências de outros equipamentos, objetos, fiação elétrica ou mesmo paredes. Dispositivos a uma mesma distância de um ponto de acesso podem receber qualidades de sinal diferentes.

Ainda segundo SIMÕES (2015), um campo de informação que pode ser recolhido pelos dispositivos móveis é o MAC Address, que permite identificar, na rede, o AP ao qual o aparelho está ligado, característica esta utilizada no desenvolvimento deste trabalho.

Como explicado por Teleco (2016), para permitir que uma estação móvel comunique-se com outras em uma rede IBSS (Independent Basic Service Set) ou um AP em uma rede infra-estrutura BSS (Basic Service Set), ela deve primeiramente encontrá-las - processo conhecido como *varredura*. Esse processo pode ser de 2 tipos: passivo, modalidade envolvendo somente a “escuta” de tráfego; e ativo (método utilizado neste trabalho), no qual a estação X executa uma varredura para extrair informações das demais estações e dos AP's, economizando tempo. Para tanto, a estação ativamente transmite *queries* (perguntas para extrair as respostas das estações numa BSS), movendo-se então para um canal e transmitindo um pacote do tipo *probe request* (requisição de sondagem): se houver BSS no canal que coincida com o SSID (Service Set Identifier) do quadro “requisição de sondagem”, a estação irá responder, enviando um quadro *probe response* (resposta de sondagem) para a estação que fez a pergunta.

O *probe request* é utilizado como meio de detecção de pessoas, pois independe

da tecnologia do tipo de estações móvel (dispositivo móvel) e está presente em todo lugar (trivial). Ainda, há muitos projetos que já fazem essa detecção através do Tshark, ou seja, há ampla documentação.

3.1.2 Modo monitor

Para que um dispositivo móvel seja identificado independente de qualquer rede ou AP, é essencial que o sistema de detecção possua uma placa e/ou adaptador de rede (NIC) que possa ser habilitado para o modo monitor.

Geralmente, uma interface de rede qualquer captura pacotes dos tipos *managed* e *beacons* que são originados por APs. Estes pacotes são transmitidos muitas vezes por segundo por APs para indicar quais redes estão realizando *broadcasting*. O modo monitor (*monitor mode*) é um modo de operação em que um NIC consegue capturar todos os tipos de pacotes sem estar associado a um AP (Acrylic WiFi, 2017) (WIRESHARK, 2017a). Dessa forma, é possível capturar todos os tipos, como os de *probe request* que são enviados de dispositivos móveis.

Neste trabalho, um Raspberry Pi com um adaptador Wi-Fi habilitado no modo monitor captura pacotes *probe request* de dispositivos móveis para que ocorra a identificação de indivíduos.

3.1.3 Tshark

O protocolo Tshark é a versão de terminal do protocolo analisador de rede Wireshark (WIRESHARK, 2017b) (WIRESHARK, 2017c). Ele é utilizado para analisar e filtrar (*sniff*) os pacotes capturados. Esse protocolo foi escolhido, pois permite realizar o estudo da rede a partir do recebimento de pacotes e seus campos, além de possuir ampla documentação, maturidade e exemplos por ser uma tecnologia aberta. Um exemplo de seu uso é o comando a seguir.

```
tshark -i wlan1 -T fields -e wlan.sa -e frame.time
```

Cada parte do comando anterior significa:

- **wlan1**: interface de rede que indica a antena Wi-Fi utilizada;
- **wlan.sa**: mostra o endereço MAC do dispositivo que enviou o pacote (em inglês, *source address*);
- **frame.time**: mostra a hora, dia e ano em que o pacote foi capturado;

A saída no terminal do comando apresentado está na Figura 10.

Figura 10 – Saída após execução de comando Tshark

80:58:f8:f3:43:9a	Jul 16, 2017 19:58:37.417872667 UTC
80:58:f8:f3:43:9a	Jul 16, 2017 19:58:37.426913501 UTC
80:58:f8:f3:43:9a	Jul 16, 2017 19:58:37.431863813 UTC
e2:11:98:a7:67:6e	Jul 16, 2017 19:59:03.309330418 UTC
e2:11:98:a7:67:6e	Jul 16, 2017 19:59:03.318259949 UTC
e2:11:98:a7:67:6e	Jul 16, 2017 19:59:03.426447189 UTC
e2:11:98:a7:67:6e	Jul 16, 2017 19:59:03.532881824 UTC
e2:11:98:a7:67:6e	Jul 16, 2017 19:59:03.645090262 UTC

Fonte: Elaborada pelas autoras.

3.1.4 Raspberry Pi

O Raspberry Pi é um computador do tamanho de um cartão de crédito que pode ser conectado a monitores, teclados, mouse e outros dispositivos. Para a detecção de dispositivos móveis um Raspberry Pi Model 3 B juntamente com um adaptador Wi-Fi são utilizados. O Raspberry foi escolhido pois oferece interface amigável de programação (terminais gráfico e de texto); possui poder de processamento para receber os milhares de pacotes, pré-processá-los e enviar para o servidor; além de entrada USB para receber uma antena Wi-Fi e também por seu tamanho pequeno (Raspberry Pi, 2017).

Outras opções foram consideradas por serem baratas, acessíveis e terem documentação aberta. Foi o caso do ESP8266 que possui um tamanho extremamente reduzido e possui o custo médio de R\$15,00 (EMBARCADOS, 2015), mas seu uso para este trabalho fica impossibilitado, pois essa tecnologia não consegue ser habilitada para o modo monitor da interface de rede (PUHL, 2016) (FERREIRA, 2016).

Uma antena Wi-Fi (Ralink MT7601U) foi equipada no Raspberry para ampliar o alcance da captura já que o propósito do sistema é detecção em zonas que podem apresentar esparsidade (espalhamento) de indivíduos e ela pôde ser habilitada para o modo monitor (subseção 3.1.2). A antena nativa do Raspberry não conseguiu ser habilitada para o *monitor mode*.

O sistema operacional utilizado com esta ferramenta foi o Kali Linux, pois, dentre os sistemas testados, foi o que suportou *drivers* para a antena Ralink utilizada. Outro ponto é que este SO foi projetado para experimentos e projetos relacionados à redes de computadores, então possui ampla documentação e exemplos nesse sentido (Kali Linux, 2017).

3.1.5 Node.js

O Node ou Node.js é um *runtime* que permite a execução Javascript fora dos navegadores, nesse caso do lado dos servidores (NODE.JS, 2017). Essa ferramenta permite a criação de APIs com entradas e saídas (I/O) não bloqueantes, ou seja, emprega programação assíncrona (DZONE, 2017). Para tal, essa tecnologia é orientada a eventos, assim que um evento ocorre funções específicas são acionadas para tratá-lo. Nesse cenário, os comandos executam em paralelo utilizando funções de *callback*. Para o sistema proposto nesse trabalho, é uma característica muito importante, já que se trata de um "sistema quase em tempo real", ou seja, é possível atender todas as requisições feitas pelo usuário para a visualização dos dados, como a postagem de arquivos provenientes do módulo sensor (??).

No GSMART, o Node é utilizado tanto na parte do servidor como na parte que executa os comandos do Tshark dentro do Kali Linux, demonstrando ainda mais sua versatilidade.

Outro componente importante do Node é a importação de pacotes produzidos pela comunidade aberta através do gerenciador de pacotes NPM (Node Package Manager), sendo essencial para facilitar e agilizar a programação. Por exemplo, o servidor deste projeto é baseado no pacote Express.js.

O Express.js é um *framework* que fornece conjunto robusto para aplicativos web, como: métodos HTTP e *middleware* (EXPRESS.JS, 2017). Na Figura 11 é possível observar o tratamento de uma das rotas de acesso à aplicação de apresentação de dados construído facilmente com o Express. Para importar este pacote, basta abrir o terminal no diretório do projeto e executar o comando `npm install --save express`.

Na linha 1, importa-se o pacote. Já na linha 11, ele faz o tratamento da rota "login", caso o usuário tente acessá-la (método GET), será retornado uma mensagem de "Você está logado!". Na linha 15, o servidor está escutando a porta 3000 para acesso. Esse trecho de código é um exemplo de como a programação com Node e seus pacotes por ser simplificada.

O último ponto para a escolha do Node.js foi o fato dele ser baseado na *engine* V8 de Javascript do Google Chrome, possibilitando utilizar Javascript tanto para o servidor quanto no lado do cliente.

3.1.6 MongoDB

O MongoDB é um banco de dados NoSQL baseado em coleções de documentos. O número de campos, conteúdo e tamanho de cada documento pode ser diferente (MONGODB, 2017). Cada um desses documentos é representado por um objeto javascript (JSON). Além disso, o MongoDB oferece API que forece *queries* rápidas e semelhantes ao SQL.

Figura 11 – Código utilizando Express.js

```
1 var express = require('express');
2 var app = express();
3
4 var myLogger = function (req, res, next) {
5   console.log('LOGGED');
6   next();
7 };
8
9 app.use(myLogger);
10
11 app.get('/login', function (req, res) {
12   res.send('Você está logado!');
13 });
14
15 app.listen(3000);
```

Fonte: Elaborada pelas autoras.

Este banco de dados foi utilizado no projeto devido a representação dos dados em JSON, já que a aplicação é inteiramente em Javascript, ou seja, facilita a manipulação pelas funções. Outro ponto é a escalabilidade, uma vez que no início do projeto o custo de uma *query* e o tamanho em *bytes* da unidade do dado não eram conhecidos. Além disso, optou-se por um banco não-relacional, pois os dados utilizados organizados em tabelas consumiriam demasiado desempenho da aplicação. O modelo de dado será apresentado no Capítulo 4.

Por fim, os documentos do MongoDB foram manipulados pelo Node.js através do pacote Mongoose que proporciona variadas funções para abstrair as *queries*.

3.2 Métodos e Etapas

A implementação deste trabalho seguiu 4 etapas. Inicialmente, foi feito o levantamento teórico acerca de *geomarketing* - devido à abrangência do tema, fez-se necessário definir qual seria o viés a ser seguido. Decidiu-se focar o desenvolvimento em conceitos relativos à contagem de pessoas e aferição de tráfego local. Foi realizado também um levantamento bibliográfico a respeito de empresas e projetos semelhantes ao proposto nesta monografia.

A segunda etapa envolveu a escolha dos materiais e tecnologias a serem utilizados na concepção de um contator de pessoas. Foram realizados testes de compatibilidade entre hardware e Sistema Operacional - no caso, utilizando-se como base o dispositivo portátil

Raspberry Pi conforme seção 3.1. Fizeram-se testes para encontrar o SO adequado às necessidades do projeto, além de testes para escolha de uma antena Wi-Fi compatível com as especificações tecnológicas, uma vez que a antena nativa do Raspberry Pi não funcionou em modo monitor, algo imprescindível na captura de pacotes de *probe request*. Após selecionar hardware e SO (Kali Linux e antena Ralink), foram escolhidas as demais tecnologias e ferramentas a serem usadas no desenvolvimento da aplicação e visualização de dados, que se daria em forma de um *webserver*: linguagens de programação, bancos de dados e servidores.

Na terceira etapa, desenvolveu-se o módulo de visualização de dados da aplicação. Primeiramente, construiu-se um servidor em Node.js e sua hospedagem num servidor em nuvem (Digital Ocean). Posteriormente, decidiu-se quais seriam as informações relevantes para serem extraídas dos dados, e construiu-se o layout: foram utilizados HTML, CSS, Javascript, EJS (Embedded JavaScript - template para apresentação de dados) e Bootstrap em forma de *dashboard* para apresentar gráficos e resumo das informações. Após isso, o MongoDB foi definido como banco de dados - utilizando-se de sua forma de organização (coleções e documentos), definiu-se o formato do dado (objetos e atributos). Por fim, os modelos de preparação e mineração de dados e o tratamento de rotas para acesso aos conteúdos do webserver (rotas de GET) e postagem de arquivos (rotas de POST).

Na última etapa, concentrou-se em desenvolver o sensor de detecção com o Raspberry Pi e a antena Ralink para adequar-se ao modelo de dado definido na etapa anterior. Desenvolveu-se um *script* que configura a antena para o modo monitor, inicia a varredura (executando o Tshark) e envia os arquivos gerados para o servidor. Nesta etapa, o programa foi aperfeiçoado aos poucos, testando a comunicação entre o Raspberry Pi e o servidor, e se a formatação dos arquivos enviados e tempo de captura estavam corretos.

De modo geral, a aplicação foi testada durante/a cada etapa - inicialmente, a partir de dados fictícios, criados no banco de dados. Posteriormente, os testes prosseguiram com os dados das inúmeras capturas realizadas.

Os principais objetivos nos testes foram:

- verificar comunicação entre módulos. Por exemplo: envio de dados do sensor para o servidor, e captura dos pacotes;
- tempo de resposta de *requests* para servidor;
- garantir programação assíncrona, que é o paradigma do Node.js;
- gerar arquivos que fossem corretamente enviados;
- garantir que as informações extraídas estivessem corretas do ponto de vista estatístico;

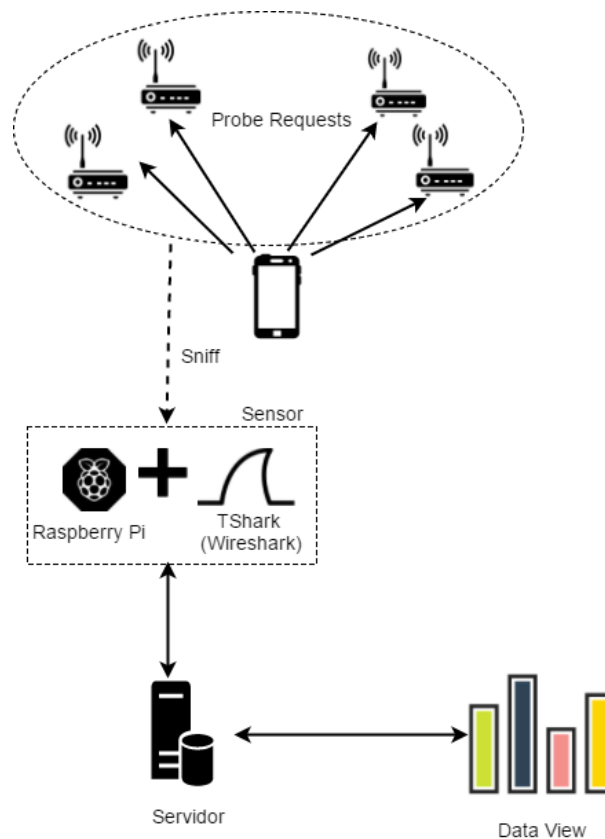
- tratar todos os tipos de erros encontrados e supostos problemas para não haver quebra no funcionamento de nenhum módulos (tratamento de exceções);
- garantir que o formato de dado manipulado fosse mantido em toda a aplicação, sem necessidade de muitas adaptações nos módulos.

Além da fase de desenvolvimento, houve a fase de testes para aferir a precisão de medição do GSMART (veja Capítulo 5).

4 ARQUITETURA DO PROJETO

Este capítulo descreve o funcionamento de cada componente do GSMART. A Figura 12 representa um esquema geral dos módulos do projeto, já a Figura 13 apresenta o mesmo esquema da visão de um diagrama de fluxo, sendo apontados entidades e processos do sistema.

Figura 12 – Arquitetura geral do sistema

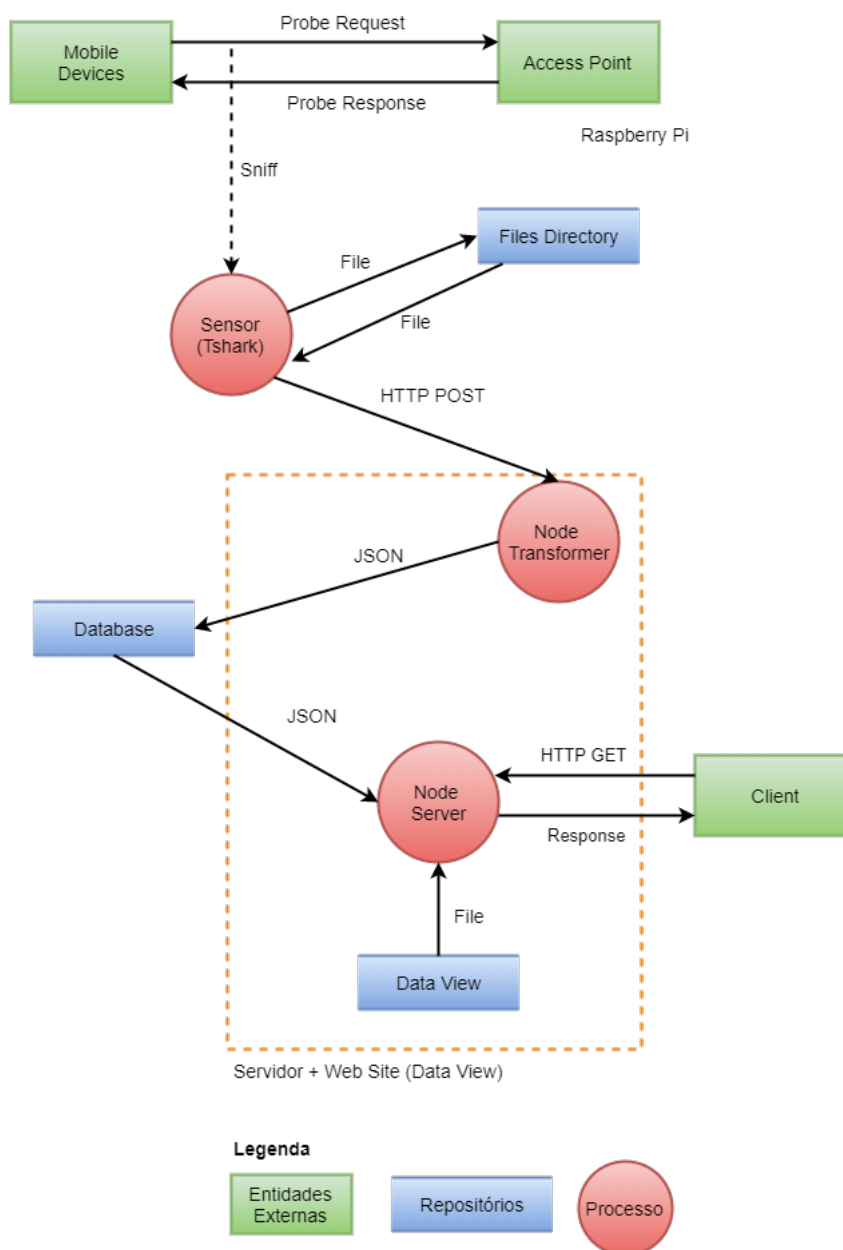


Fonte: Elaborada pelas autoras.

4.1 Sensor

O módulo Sensor é um processo que roda num Raspberry Pi Model 3 B equipado de uma antena Ralink no modo monitor. Este processo é um programa escrito em Node.js que: habilita o modo monitor da antena, executa comandos Tshark para capturar de pacotes *probe request*, salva os endereços MAC capturados num arquivo que, posteriormente, é recuperado do diretório de arquivos (*Files Directory*) e enviado ao servidor.

Figura 13 – Diagrama de Fluxo



Fonte: Elaborada pelas autoras.

A Figura 14 contém uma parte do código que executa dentro do sensor. Ele funciona em 3 fluxos:

- **startDate**: esse fluxo, representado na linha 78 do código, executa a função "start-Date" cada 60 minutos, criando um arquivo com o nome "ZONA_YYYY-MM-DD_HH:00", especificando zona, dia e hora de detecção. Logo após, salva o arquivo e inicia o fluxo a seguir;
- **scan**: este fluxo (linha 78) repete a função "scan" a cada 40 segundos durante 50

Figura 14 – Trecho de código do sensor

```
64 function scan(){
65     exec(`tshark -i wlan0 -a duration:40 -Y "wlan.fc.type_subtype eq 4"
66         -T fields -e wlan.sa > temp`)
67     .then((result) => {
68         countScan++;
69         fs.readFile('temp', "utf8", (err, data) =>{
70             fs.appendFile(fileName,data, (err) => {
71                 if(countScan == 50)
72                     sendNewFile();
73             });
74         });
75     });
76 }
77
78 Repeat(startDate).every(60, 'minutes').for(300, 'minutes').start.now();
79
80 Repeat(sendAllFiles).every(5, 'minutes').for(300, 'minutes').start.now();
```

Fonte: Elaborada pelas autoras.

minutos que executa o comando Tshark (veja subseção 3.1.3) que vai armazenando os endereço MAC num arquivo temporário. Quando as execuções do Tshark chegam na contagem 50 o conteúdo do arquivo é copiado para o arquivo gerado no fluxo anterior, então é enviado ao servidor;

- **sendAllFiles**: este fluxo (linha 80) repete a função "sendAllFiles" a cada 5 minutos enviando todos os arquivos que tiveram falha no seu primeiro envio. capturado.

4.2 Node Transformer

Após receber o arquivo, o servidor salva-o no diretório e encaminha os endereços MAC capturados para o módulo de transformação.

O Node Transformer é um módulo dentro do *webserver* que transforma os endereços MAC em dados formatados. de dados. Primeiro, são eliminadas os endereços repetidos. Depois, a estrutura de dado (objeto) que representa um registro do banco é preenchida com a informações. Uma parte importante dessa preparação é que o fabricante de cada dispositivo é obtido através de uma API¹ de consulta de endereços MAC. Por fim, o objeto preenchido é salvo no banco de dados.

¹ <<https://macvendors.com/>>

4.2.1 Formatação do dado

A formatação do dado foi pensada de acordo com a unidade de tratamento escolhida para a medição tráfego, o dia. Dentro da aplicação e do banco, o dado possui o formato do JSON da Figura 15.

Figura 15 – Formatação do dado

```
1  {
2      'day': '2017-11-25',
3      'hour': '17:00',
4      'zone': 'LTIA',
5      'macs': [
6          {
7              'customer': false,
8              'vendor': 'Samsung Electronics Co.,Ltd',
9              'mac': '98:39:8e:a0:d7:d3'
10         },
11         {
12             'customer': true,
13             'vendor': 'AzureWave Technology Inc.',
14             'mac': '6c:ad:f8:dc:05:a2'
15         }
16     ]
17 }
```

Fonte: Elaborada pelas autoras.

Cada objeto como o da Figura 15, possui os seguintes campos sobre a medição:

- **day**: dia;
- **hour**: hora;
- **zone**: zona em que foi realizada;
- **macs**: vetor de todos os endereços MAC capturados;
- **customer**: se o endereço captado já foi capturado outro(s) dia(s);
- **vendor**: fabricante do dispositivo;
- **mac**: endereços MAC.

O maior ganho com essa formatação é durante as *queries* e processamento dos dados. O dado pode ser recuperado por zona ou por dia, dependendo da visualização de dados selecionada.

4.3 Node Server

Este módulo que engloba o módulo da seção 4.2, também: define quais recursos recursos da aplicação são acessadas pelo cliente e como os dados são processados segundo cada visualização de dados.

4.3.1 Deliberando recursos

O Node Server é um *web service* baseado no padrão REST (Representational State Transfer, em português Transferência de Estado Representacional). Adotando esse padrão implica que temos uma RESTful API (Jackson Pires, 2017). De modo geral, a aplicação define URIs que através de verbos HTTP (GET e POST), fornecem à requisição do cliente algum recurso específico. Por exemplo, ao acessar a URL do servidor acompanhado de `"/statistics/LTIA"`, a API desenvolvida responde ao cliente com uma página HTML com informações referentes a zona LTIA.

Todas as URIs possíveis foram definidas através de um *middleware* com o pacote do Node, o Express.js². Este emprega rotas para definir quais recursos serão fornecidos ao cliente. Na Figura 16 temos o exemplo de trecho de código do *middleware* da aplicação. Nesse exemplo, assim que o servidor recebe um requisição HTTP GET, após um processamento de dados, responde ao cliente com a visualização de dados “statistics” com o dado em forma de JSON (linha 5). Já Figura 17 descreve todas as rotas (URIs) acessíveis ao cliente.

Figura 16 – *Middleware* com Express.js

```
1  app.get('/statistics/:zone/:year/:month/:day', function(req, res) {  
2      .  
3      .  
4      .  
5      res.render('statistics', {obj: obj})  
6  });
```

Fonte: Elaborada pelas autoras.

² <<http://expressjs.com/>>

Figura 17 – Rotas do Express

URI	Método HTTP	Corpo da requisição	Resultado
/upload	POST	Vazio	Salva dados do arquivo no banco de dados.
/	GET	Vazio	Renderiza página HTML para escolha da zona.
/comparing/:zone	GET	Vazio	Renderiza página HTML para tela de comparação segundo a zona escolhida.
/statistics/:zone	GET	Vazio	Renderiza HTML para o dashboard com os gráficos da zona escolhida.
/statistics/:zone/:year/:month/:day	GET	Vazio	Renderiza HTML para o dashboard com os gráficos da zona e dia escolhidos.

Fonte: Elaborada pelas autoras.

4.3.2 Processando dados

Além de deliberar recursos, o Node Serve recupera os dados e processá-os para apresentar ao usuário. Assim que o cliente faz a requisição ao servidor, dados referentes à zona e/ou dia selecionados são recuperados do banco de dados. Em seguida, esses elementos passam por outros módulos que retornam JSON com as informações necessárias para apresentar ao usuários.

4.4 Visualização de Dados

5 RESULTADOS E DISCUSSÃO

Essa seção descreve o procedimento de testes do sistema e os resultados obtidos a partir dos mesmos.

Durante e após o desenvolvimento e interligação de tecnologias e ferramentas utilizadas conforme seção 3.2, executaram-se vários testes a fim de validar se a captura de pacotes e contagem estavam ocorrendo como previsto. Ao longo dessas verificações, foram feitos também ajustes a fim de otimizar a precisão dos resultados, levando-se em conta o fato de que a aferição é uma estimativa média de tráfego e que existem limitações tecnológicas envolvidas. Fizeram-se testes extensivos (tanto em ambiente controlado quanto comercial) a fim de garantir-se a eficácia do dispositivo e também criar um banco de dados real, fornecendo os dados necessários ao *website* para processamento e geração de informação.

De modo geral, a aplicação foi testada durante cada etapa de desenvolvimento (veja seção 3.2), inicialmente a partir de dados fictícios, criados no banco de dados. Posteriormente, os testes seguiram com os dados das inúmeras capturas realizadas.

5.1 Testes e validação do projeto

Para a validação do sistema proposto, executaram-se testes unitários, de integração e validação. Os principais objetivos dos testes foram:

Os testes também tiveram objetivo de verificar se o volume de dados e/ou o tempo de captura interfeririam na eficiência do contador, já que foram realizadas verificações de tempo curtas (em torno de 5 a 10 minutos) e longas (3 horas ou mais ininterruptas).

Os testes unitários e de integração do sistema resumem-se em:

- detecção dos dispositivos móveis;
- comunicação entre Raspberry Pi e servidor;
- comunicação servidor e interface;
- processamento de dados capturados em dados desejados;
- determinar se sistema consegue contar pessoas.

Já a validação focou-se em determinar se o sistema consegue ou não realizar a contagem e o tráfego de pessoas. Para tanto, foram feitos testes em ambientes controlados

e não controlados: os controlados são zonas pequenas, em que se conhece previamente o número de pessoas, sendo possível conferir o resultado da detecção. Nos ambientes não-controlados, a medição do tráfego é feita sem que se saiba quantas pessoas poderão adentrar a zona de teste. A taxa de confiabilidade no sistema é baseada no desvio padrão dos testes realizados em ambiente controlado - o sistema final é considerado aplicável ou não conforme o desvio padrão determinado (dependendo se este encontra-se ou não dentro dos limites estabelecidos).

O procedimento de testes ocorreu da seguinte maneira: inicialmente, foram feitos em ambiente controlado, a fim de observar o comportamento do sistema desenvolvido. Após todas as verificações controladas, testou-se o sistema em ambiente não-controlado, ou seja: um estabelecimento comercial - nessa etapa, o projeto verificou o desempenho do sistema em ambiente real, com grande quantidade de dispositivos móveis.

Como ambiente controlado, foram escolhidos o LTIA (Laboratório de Tecnologia da Informação Aplicada), localizado na UNESP de Bauru/SP (contando-se com as devidas autorizações dos responsáveis) e também a residência das autoras deste trabalho. Percebeu-se que, nos testes realizados na zona do LTIA, a contagem ficou um pouco acima do esperado devido à grande quantidade de máquinas e dispositivos que se encontravam no recinto, consideravelmente superior à de pessoas. Nos testes realizados na moradia das autoras, o tráfego obtido mostrou-se bem mais próximo ao esperado, considerando-se houve a captura de pacotes originados em apartamentos adjacentes, já que ambas residem em condomínio.

Para verificação em ambiente não-controlado, realizaram-se testes no estabelecimento comercial Camflam Lanches, em Bauru/SP (também contando-se com as devidas autorizações). A captura de pacotes e geração de dados deu-se normalmente, e os resultados obtidos mostraram-se dentro do previsto. Neste ambiente, foram feitos somente testes longos, sem interrupção, com objetivo de aproximar-se ao máximo do que seria o ambiente real no qual o dispositivo de fato operaria.

5.2 Análise de Riscos

Alguns fatores e requisitos funcionais estão sujeitos a desvios de planos ou falhas técnicas, sendo necessário implementar um plano de contingência para garantir a acurácia do sistema, conforme descrito abaixo:

- Falha da detecção de dispositivos (precisão): foram feitas duas formas de detecção através do protocolo Tshark - ambas garantem que os dados de uma e outra são verídicos - caso uma forma de detecção falhe, há outra para detectar os dispositivos móveis;

- Processamento em servidor: o desenvolvimento do processamento de dados no servidor é complexo, considerando-se que trabalhamos com estatísticas. Optou-se então por uma *cloud*, a fim de obter-se melhor desempenho e segurança possíveis;
- Falhas de conexão: O Raspberry Pi está sujeito a perder a conexão com a rede e ficar impossibilitado de repassar os dados ao servidor - logo, um backup é armazenado dentro do aparelho: no caso de falha na conexão, quando a mesma retornar, os arquivos serão enviados ao servidor. Existe ainda a possibilidade do uso de um Modem 3G para garantir o envio.

6 CONCLUSÃO

O foco principal deste trabalho foi disponibilizar um sistema *opensource* que aplique os conceitos de *geomarketing* vistos, abrangendo o estudo de redes de computadores (redes sem fio) e dispositivos móveis aplicados à administração, com ênfase em gestão da informação. A partir conceitos de *geomarketing* provenientes de autores, consultorias e trabalhos semelhantes, foi possível agrupar a contagem de pessoas por períodos e descobrir informações relevantes para a tomada de decisão estratégica em negócios (subseção 2.1.2).

O protótipo desenvolvido consegue realizar a contagem de pessoas em zonas específicas, porém ressaltamos que, em áreas muito grandes, faz-se necessário multiplicar o número de sensores de acordo com a cobertura necessitada, a fim de manter-se resultados pontuais. Foi ainda observado que, em locais com alta densidade de máquinas e dispositivos fixos por pessoa, como o LTIA (Laboratório de Tecnologia da Informação Aplicada), a eficácia da estimativa foi comprometida: como a quantidade de dispositivos é muito mais alta do que a de pessoas, esse fator pode inflar as estimativas, exibindo uma frequência de pessoas acima do real. Logo, por se basear no *MAC Address* de dispositivos para contagem, o sistema não é indicado para medição em locais com alta concentração de aparelhos fixos. O contador se comporta com eficácia em áreas onde existe alta mobilidade (grande entrada e saída de pessoas), como restaurantes e lojas, onde os transeuntes geralmente portam um dispositivo móvel pessoal (*smarthphone* ou *tablet*).

Outro ponto a ser salientado é que, como as soluções de *geomarketing* são, em maior parte, de âmbito privado, grande parte da fundamentação teórica volta-se para consultorias, estudos de caso e trabalhos correlatos. Por essa razão, foi difícil também o levantamento de dados financeiros para estimarem-se os custos do serviço. Em livros, muitos conceitos teóricos podem ser encontrados - entretanto, não como são aplicados, existiram dificuldades no desenvolvimento do sistema proposto, como decidir de que forma organizar a fase estatística na agregação/ contagem de pessoas e processamento de dados: a precisão do sistema e as informações coletadas devem ser relevantes e numerosas (generalização).

REFERÊNCIAS

Acrylic WiFi. *Monitor Mode and Native Capture Mode in Acrylic Wi-Fi*. 2017. Disponível em: <<https://www.acrylicwifi.com/en/blog/wifi-monitor-mode/>>.

ARAGÃO, P. S. S. de. Geomarketing: Modelos e Sistemas, com Aplicações em Telefonía. 2005. Disponível em: <<http://www.lis.ic.unicamp.br/wp-content/uploads/2014/09/aragao.pdf>>.

AXPER. *3D People Counters for retail | Axper*. 2017. Disponível em: <http://axper.com/people/{_}count>.

CISCO. *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper*. 2017. Disponível em: <<http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html>>.

Cisco Blogs. *How Many Internet Connections are in the World? Right. Now*. 2013. Disponível em: <<https://blogs.cisco.com/news/cisco-connections-counter>>.

CLIQUET, G. *Geomarketing : Methods and Strategies in Special Marketing*. 1. ed. [S.l.]: Lavosier/ISTE, 2006. 327 p. ISBN 190520907X.

DENSITY. *Density | A People Counter & API*. 2017. Disponível em: <https://www.density.io/{_}>.

DUARTE, V. *O que é Geomarketing?* 2016. Disponível em: <<https://blog.geofusion.com.br/infografico-o-que-e-geomarketing>>.

DZONE. *What are the Benefits of Node.js?* 2017. Disponível em: <<https://dzone.com/articles/what-are-benefits-nodejs>>.

EMBARCADOS. *Apresentando o módulo ESP8266*. 2015. Disponível em: <<https://www.embarcados.com.br/modulo-esp8266/>>.

EVANS, D. The Internet of Things How the Next Evolution of the Internet Is Changing Everything. *Cisco IBSG*, 2011. Disponível em: <http://www.cisco.com/c/dam/en/{_}us/about/ac79/docs/innov/IoT/{_}IBSG/{_}0411>.

EXPRESS.JS. *Express - framework de aplicativo da web Node.js*. 2017. Disponível em: <<http://expressjs.com/pt-br/>>.

FERREIRA, L. C. P. *Sistema localizador interior de baixo custo*. 79 p. Tese (Mestrado) — Instituto Superior de Engenharia de Lisboa, 2016. Disponível em: <<http://repositorio.ipl.pt/handle/10400.21/6162>>.

FIGUERA, C. et al. Time-Space Sampling and Mobile Device Calibration for WiFi Indoor Location Systems. *IEEE Transactions on Mobile Computing*, v. 10, n. 7, p. 913–926, jul 2011. ISSN 1536-1233. Disponível em: <<http://ieeexplore.ieee.org/document/5765967/>>.

GARTNER. *Gartner Says the Internet of Things Will Transform the Data Center*. 2014. Disponível em: <<http://www.gartner.com/newsroom/id/2684616>>.

HOU, Y.-L.; PANG, G. K. H. People Counting and Human Detection in a Challenging Situation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, v. 41, n. 1, p. 24–33, jan 2011. ISSN 1083-4427. Disponível em: <<http://ieeexplore.ieee.org/document/5580105/>>.

Ipsos Retail Performance. *How to Choose the Right People Counter*. 2015. Disponível em: <<https://www.ipsos-retailperformance.com/resources/blog/how-to-choose-right-people-counter/>>.

Jackson Pires. *O que é API? REST e RESTful?* 2017. Disponível em: <<https://becode.com.br/o-que-e-api-rest-e-restful/>>.

Kali Linux. *Kali Linux Official Documentation*. 2017. Disponível em: <<https://www.kali.org/kali-linux-documentation/>>.

KIM, J. et al. Real-time vision-based people counting system for the security door. ... *Technical Conference on Circuits Systems*, ... , v. 3672, n. July 2016, 2002. Disponível em: <<http://scholar.google.com/scholar?hl=en{%&}btnG=Search{%&}q=intitle:Real-time+Vision-based+People+Counting+System+for+the+Security+>>>.

KOTLER, P.; KELLER, K. *Administração de Marketing*. 14. ed. [S.l.]: Pearson, 2013. 796 p. ISBN 9788581430003.

Libelium Comunicaciones Distribuidas S.L. *Meshlium X treme Technical Guide*. [S.l.]: Libelium, 2017. 188 p.

LUO, J. et al. Real-time people counting for indoor scenes. *Signal Processing*, v. 124, p. 27–35, 2016. ISSN 01651684. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0165168415003801>>.

MA, H.; ZENG, C.; LING, C. X. A Reliable People Counting System via Multiple Cameras. *ACM Transactions on Intelligent Systems and Technology*, v. 3, n. 2, p. 1–22, feb 2012. ISSN 21576904. Disponível em: <<http://dl.acm.org/citation.cfm?doid=2089094.2089107>>.

MANGINI, E. R.; LUZ, L. F. D.; CONEJERO, M. A. Modelo de Análise de Localização e Aplicações de Geomarketing em Transporte Público de Alta Capacidade: o Caso da Companhia Paulista de Trens Metropolitanos. 2014. Disponível em: <<http://www.aedb.br/seget/arquivos/artigos14/6120100.pdf>>.

MONGODB. *MongoDB 3.6*. 2017. Disponível em: <<https://www.mongodb.com/mongodb-3.6?jmp=homepage{%&}jmp=homep>>.

NODE.JS. *Node.js*. 2017. Disponível em: <<https://nodejs.org/>>.

PUHL, L. *Habilitando um prédio a localizar contextualmente dispositivos utilizando redes sem fio*. 85 p. Tese (Graduação) — Universidade Estadual Paulista Júlio de Mesquita Filho, 2016.

Raspberry Pi. *Raspberry Pi 3 Model B*. 2017. Disponível em: <<https://www.raspberrypi.org/products/raspberry-pi-3-model-b/>>.

SCHOLLZ. *howmanypeoplearearound*. 2017. Disponível em: <<https://github.com/schollz/howmanypeoplearearound/>>.

SEABRA, A. L. D. C. Geomarketing: Estudo De Áreas De Influência De Um Restaurante De Healthy Food Em Natal. p. 105, 2014.

SEED. *Contagem e fluxo de Pessoas*. 2017. Disponível em: <http://www.seeddigital.com.br/o-que-fazemos/{\#}servicos{_}de>.

SHOPPERTRAK. *Análise de Perímetro*. 2017. Disponível em: <<https://br.shoppertrak.com/solutions/analysis-de-perimetro/>>.

SIMÕES, D. M. Navegação indoor baseada na rede WiFi como suporte a serviços baseados na localização: Estudo de caso no campus da UL. *Mestrado em Sistemas de Informação Geográfica*, 2015.

TELECO. *teleco.com.br*. 2008. Disponível em: <http://www.teleco.com.br/tutoriais/tutorialww1/pagina{_}4>.

TELECO. *teleco.com.br*. 2016. Disponível em: <http://www.teleco.com.br/tutoriais/tutorialrwan2/pagina{_}6>.

TELECO. *teleco.com.br*. 2017. Disponível em: <<http://www.teleco.com.br/ncel.asp>>.

TRAF-SYS. *People Counters, Retail Traffic Counting, and Pedestrian Door Counters*. 2017. Disponível em: <<http://www.trafsys.com/people-counting/>>.

V-COUNT. *People Counter, People Counters, visitor counter, Customer Counter, Footfall, Door Counter, Retail Traffic Counter*. 2017. Disponível em: <<http://v-count.com/>>.

VENKATESH ANKAN BANSAL, K. S. People Counting in High Density Crowds from Still Images. *International Journal of Computer and Electrical Engineering*, v. 7, n. 5, p. 316–324, 2015. ISSN 17938163. Disponível em: <<http://www.ijcee.org/index.php?m=content{&}c=index{&}a=show{&}catid=7>>.

WADSWORTH, C. *Measuring Retail Store Traffic: How People Counting Works*. 2013. Disponível em: <<http://www.trafsys.com/what-counting-can-do-for-profits-part-4-of-4/>>.

WIRESHARK. *CaptureSetup/WLAN*. 2017. Disponível em: <<https://wiki.wireshark.org/CaptureSetup/WLAN>>.

WIRESHARK. *tshark\ -\ The\ Wireshark\ Network\ Analyzer\ 2.0.0*. 2017. Disponível em: <<https://www.wireshark.org/docs/man-pages/tshark.html>>.

WIRESHARK. *Wireshark · Go Deep*. 2017. Disponível em: <<https://www.wireshark.org/>>.

Zebra Technologies. *Cidade Jardim Creates Personalized Shopping Experience with Zebra Wireless Solution*. 2016. Disponível em: <<https://www.zebra.com/us/en/about-zebra/newsroom/press-releases/2016/cidade-jardim-creates-personalized-shopping-experience-with-zebra-wireless-solution.html>>.