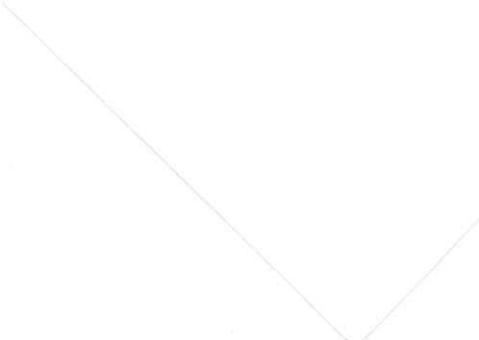


Universidade Federal de Campina Grande  
Centro de Engenharia Elétrica e Informática  
Coordenação de Pós-Graduação em Ciência da Computação

Dissertação de Mestrado



Seleção de Anúncios em Sinalização Digital  
baseada na Segmentação de Mercado e  
Comportamento do Consumidor

Leonardo Soares e Silva

Campina Grande, Paraíba, Brasil  
Fevereiro – 2012

Universidade Federal de Campina Grande  
Centro de Engenharia Elétrica e Informática  
Coordenação de Pós-Graduação em Ciência da Computação

Seleção de Anúncios em Sinalização Digital baseada  
na Segmentação de Mercado e Comportamento do  
Consumidor

Leonardo Soares e Silva

Dissertação submetida à Coordenação do Curso de Pós-Graduação em  
Ciência da Computação da Universidade Federal de Campina Grande -  
Campus I como parte dos requisitos necessários para obtenção do grau  
de Mestre em Ciência da Computação.

Área de Concentração: Ciência da Computação

Linha de Pesquisa: Computação pervasiva

Hyggo Oliveira de Almeida e Angelo Perkusich

(Orientadores)

Campina Grande, Paraíba, Brasil

©Leonardo Soares e Silva, 05/02/2012

**DIGITALIZAÇÃO:**  
**SISTEMOTECA - UFCG**

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG**

S586s Silva, Leonardo Soares e.  
Seleção de anúncios em sinalização digital baseada na segmentação de mercado e comportamento do consumidor / Leonardo Soares e Silva. – Campina Grande, 2012.  
83 f. : il.

Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática.

Orientadores: Prof. Dr. Hyggo Oliveira de Almeida, Prof. Dr. Angelo Perkusich.  
Referências.

1. Computação Pervasiva. 2. Sinalização Digital. 3. Publicidade Pervasiva. 4. Comportamento do Consumidor. 5. Sistema de Recomendação. I. Título.

CDU 004.4'2:64.033(043)

**"SELEÇÃO DE ANÚNCIOS EM SINALIZAÇÃO DIGITAL BASEADA NA  
SEGMENTAÇÃO DE MERCADO E COMPORTAMENTO DO CONSUMIDOR"**

**LEONARDO SOARES E SILVA**

**DISSERTAÇÃO APROVADA EM 29/02/2012**

  
**HYGGO OLIVEIRA DE ALMEIDA, D.Sc**  
**Orientador(a)**

  
**ANGELO PERKUSICH, D.Sc**  
**Orientador(a)**

  
**LEANDRO DIAS DA SILVA, D.Sc**  
**Examinador(a)**

  
**KYLLER COSTA GORGÔNIO, Dr.**  
**Examinador(a)**

**CAMPINA GRANDE - PB**

## Resumo

Em uma sociedade cada vez mais individualista e exigente, o aumento na demanda por produtos personalizados e adaptados às necessidades individuais se reflete diretamente na publicidade. Consumidores estão expostos diariamente a inúmeros meios de comunicação e requerem anúncios de produtos de seu interesse para que a mensagem seja lembrada.

Uma alternativa para tornar o anúncio relevante é adequá-lo às expectativas do consumidor. A publicidade pervasiva se mostra viável à realização desse processo, por possibilitar a veiculação de anúncios adaptados ao contexto do consumidor.

A sinalização digital é um exemplo de publicidade pervasiva e vem ganhando destaque pelo seu crescimento, oportunidades apresentadas e pelo barateamento das tecnologias que tornam viável a sua utilização. A sinalização digital possibilita a entrega de anúncios adaptados às necessidades dos consumidores, utilizando elementos gráficos com maior potencial para atrair sua atenção, além de possibilitar a exibição dos anúncios em momentos oportunos.

Com o objetivo de tornar a sinalização digital mais adaptada aos consumidores, diversas abordagens foram propostas. No entanto, a maioria dessas abordagens depende de interações explícitas com o consumidor, o que pode não ocorrer devido à casualidade desse meio de comunicação.

Neste trabalho propõe-se um método para tornar mais efetiva a seleção de anúncios sem a necessidade de interação e retorno do consumidor. O método proposto baseia-se no uso da informação sobre o público alvo e interesses relacionados aos produtos.

O método foi testado e validado com a comparação a outras abordagens identificadas na literatura e de propósitos similares. Os testes foram realizados em um experimento com a participação de 112 entrevistados, obtendo-se resultados satisfatórios e compatíveis com abordagens que exigem a interação e o retorno explícito do consumidor.

## **Abstract**

In a world where the consumer's individual demand is ever growing, the increase in the desire for custom products causes a direct impact in publicity. Consumers are daily exposed to a myriad of communications media and require advertisements for products of interest.

An alternative to make advertisement relevant to the interests of the consumers is making it customizable. Pervasive advertising makes feasible that process, because it allows advertisements to be adapted to the consumer's context.

Digital Signage is an example of pervasive advertising and has received increasing attention for its growth in adoption. Price drop on related technologies makes its adoption viable. Digital signage allows adapting and presenting advertisements to the consumers, using graphical elements which can be used to get further attention from the consumers. Also, this media can make the advertisements situational.

Several approaches were proposed in order to adapt Digital Signage for the consumers. Most of these, however, depend on the explicit interaction with the consumer which may not take place due to the casualness of the communication media.

In this work, we propose a method for making the advertisements selection more effective without the need for interaction with the customers. This method is based on information about the target audience and products of interest.

The method was validated, tested and compared with other methods with similar purposes. The tests were performed in an experiment involving 112 interviewed. The results were satisfactory and compatible with approaches that require explicit interaction.

## **Agradecimentos**

Primeiramente, agradeço à Deus que me ilumina todos os dias.

Aos meus familiares, pelo apoio em momentos difíceis e por estarem presentes em todos os momentos.

Agradeço em especial à minha namorada, Aline, por compreender e me apoiar nas situações onde não pude estar presente, e também pelas palavras de motivação que me fortaleceram nos momentos em que precisei.

Aos meus orientadores, Prof. Dr. Hyggo Almeida e Prof. Dr. Angelo Perkusich, meu reconhecimento pelo apoio e orientação que foram fundamentais às decisões tomadas. Agradeço pela oportunidade em poder trabalhar com vocês.

Ao Prof. Ms.c Fred Bublitz, pelos conselhos e sugestões que agregaram grande valor à este trabalho.

Aos colegas de mestrado pelo convívio e amizade.

Aos professores de mestrado pelo conhecimento provido durante as disciplinas.

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Problemática . . . . .	5
1.2	Objetivos . . . . .	7
1.2.1	Objetivo geral . . . . .	7
1.2.2	Objetivos específicos . . . . .	7
1.3	Relevância . . . . .	8
<b>2</b>	<b>Fundamentação teórica</b>	<b>10</b>
2.1	Computação pervasiva . . . . .	10
2.2	Publicidade pervasiva . . . . .	11
2.3	Sinalização digital . . . . .	13
2.4	Anúncios publicitários . . . . .	16
2.4.1	Relevancia dos anúncios publicitários . . . . .	17
2.4.2	Teoria sobre o comportamento do consumidor . . . . .	18
2.4.3	Segmentação de mercado e público-alvo . . . . .	19
2.5	Lógica nebulosa . . . . .	21
2.5.1	Propriedades . . . . .	22
<b>3</b>	<b>Trabalhos relacionados</b>	<b>28</b>
3.1	MobiDIC . . . . .	28
3.2	BluScreen . . . . .	30
3.3	Sistema de recomendação para painéis públicos . . . . .	33
3.4	Targeted advertisement in retail environments . . . . .	35
3.5	Advertisement Recommendation Algorithm (ARA) . . . . .	37

3.6	MyAds . . . . .	39
3.7	Comparação entre os métodos . . . . .	41
<b>4</b>	<b>Método para seleção de anúncios em sinalização digital</b>	<b>44</b>
4.1	Introdução . . . . .	44
4.2	Modelo para representação dos anúncios e estabelecimentos . . . . .	46
4.3	Método de recomendação baseado na similaridade de interesses . . . . .	49
4.4	Método de seleção baseado na proximidade entre o público-alvo dos anúncios e do estabelecimento . . . . .	51
4.5	Hibridização dos métodos de seleção de anúncios . . . . .	56
4.6	Considerações sobre os métodos . . . . .	59
<b>5</b>	<b>Validação empírica</b>	<b>61</b>
5.1	Introdução . . . . .	61
5.2	Seleção das variáveis . . . . .	62
5.3	Condução do experimento . . . . .	62
5.3.1	Classificação dos anúncios e estabelecimentos . . . . .	63
5.3.2	Métrica para avaliação dos métodos . . . . .	64
5.4	Primeira etapa . . . . .	65
5.5	Segunda etapa . . . . .	66
5.6	Resultados . . . . .	67
5.6.1	Inferência do perfil do entrevistado a partir do público-alvo e categorias de interesse do estabelecimento . . . . .	67
5.6.2	Comparação entre os métodos . . . . .	70
5.7	Análise e interpretação dos resultados . . . . .	72
5.7.1	Inferência do perfil do entrevistado . . . . .	72
5.7.2	Análise da inferência do gênero . . . . .	73
5.7.3	Análise da inferência da idade . . . . .	74
5.7.4	Análise da inferência da classe econômica . . . . .	74
5.7.5	Avaliação dos métodos . . . . .	74
5.8	Conclusão dos resultados . . . . .	75



## Lista de Figuras

1.1	Exemplo de sinalização digital. . . . .	2
1.2	Exemplo de publicidade interna. . . . .	3
1.3	Exemplo de publicidade ao ar livre. . . . .	3
1.4	Exemplo de sistema de sinalização digital instalado em um banheiro. . . . .	4
2.1	Sinalização digital instalada em ônibus [28]. . . . .	14
2.2	Processo de criação, distribuição e instalação de anúncios publicitários (Ciclo CDI) [14]. . . . .	15
2.3	Anúncio do veículo Land Rover Freelander. . . . .	23
2.4	Diagrama de Venn para um conjunto nebuloso. . . . .	23
2.5	Funções de pertinência. [4]. . . . .	24
3.1	Sinalização digital utilizada no estudo e exemplo do cupom exibido [28]. . . . .	29
3.2	Arquitetura da abordagem BluScreen [19]. . . . .	31
3.3	Resultados do primeiro experimento. No eixo y está a quantidade de anúncios exibidos e no eixo x a variação na quantidade de dispositivos [39]. . . . .	32
3.4	Resultados do segundo experimento relacionados a relevância dos anúncios ( <i>relevant</i> ), capacidade da audiência em lembrar dos anúncios exibidos ( <i>recall</i> ) e anúncios lembrados, mas que não foram exibidos ( <i>extra</i> ) [19]. . . . .	33
3.5	a) exemplo de aplicação do sistema; b) arquitetura do sistema [39]. . . . .	36
3.6	Arquitetura do sistema de sinalização digital [46]. . . . .	38
3.7	Arquitetura MyAds [9]. . . . .	40
3.8	Resultados do experimento realizado para o primeiro cenário [9]. . . . .	42
3.9	Resultados do experimento realizado para o segundo cenário [9]. . . . .	42

---

4.1	Anúncios de pacotes para o carnaval. . . . .	53
4.2	Cenário de utilização do modelo. . . . .	58
5.1	Percentuais para o acerto e quantidade de erros para a inferência dos interesses do entrevistado. . . . .	68
5.2	Percentuais para o acerto e a diferença em anos entre a idade inferida e a idade real do entrevistado. . . . .	69
5.3	Percentuais de acerto e erro para a inferência do gênero. . . . .	69
5.4	Percentuais de acerto e erro para a inferência do gênero. . . . .	70
5.5	Gráfico com os percentuais da precisão, taxa de erro e dos anúncios avaliados como pouco interessantes. Legenda: alt - Aleatorio, M.S - Método de seleção pela similaridade dos interesses, M. P.A. - Método de seleção pela proximidade entre os públicos-alvos, M.H - Método híbrido, F.C. P. - Filtragem colaborativa por correlação de <i>Pearson</i> , F.C. S. - Filtragem colaborativa por <i>Slope One</i> . . . . .	72

## Lista de Tabelas

2.1	Variáveis usadas no processo de segmentação de mercado. . . . .	20
3.1	Tabela comparativa dos métodos. . . . .	43
4.1	Variáveis usadas no processo de segmentação de mercado. . . . .	54
5.1	Precisão das abordagens. . . . .	71
5.2	Taxa de erro das abordagens. . . . .	71

# Capítulo 1

## Introdução

A área de marketing é responsável pela entrega de valores aos clientes através da concepção de produtos e serviços adequados às suas necessidades e interesses [24]. Ela é composta por um conjunto de processos que são direcionados a servir um determinado mercado. Parte desse processo envolve a comunicação que a empresa deve adotar para potencializar sua relação com os consumidores e a publicidade é um dos elementos pertencentes a esse processo que tornam isso viável.

O setor de publicidade é o responsável pela divulgação de ideias, conceitos e valores praticados por uma empresa. Essa divulgação é concretizada através dos anúncios publicitários. De acordo com Kotler [30] um anúncio é um processo de comunicação, pago, para promover ideias, produtos ou serviços, a ser realizado para um público-alvo em específico, durante um período de tempo específico.

O objetivo maior da publicidade é criar novas estratégias para informar, lembrar ou persuadir o consumidor para a escolha do produto anunciado. Por ser direcionada ao consumidor, a publicidade está em constante evolução visando adaptar-se às modificações de hábitos e comportamentos da sociedade. Dentre essas transformações, destacam-se as recentes mudanças no perfil do consumidor, que está mais exigente e individualista, provocando modificações nas campanhas publicitárias e elevando a demanda por conteúdo mais direcionado e personalizado [8].

Para atender a esse novo perfil de consumidor, as empresas de publicidade estão à procura de mais informações sobre seu público-alvo e novas formas de comunicação para alcançar a audiência com mensagens mais adaptadas aos seus interesses. Os meios de comunicação

tradicionais, como televisão, rádio e jornal, são afetados diretamente por essas mudanças. Pesquisas recentes mostram um declínio, nunca ocorrido, no número de domicílios com televisão nos Estados Unidos<sup>1</sup>. Em paralelo há uma redução no número de casas americanas com televisão por assinatura<sup>2</sup>. A redução no número de leitores de jornais impresso é outro reflexo dessas mudanças. Dados mostram que a publicação de notícias online no ano passado superou em quantidade, pela primeira vez, os jornais impressos. Há também a perspectiva de que em 2013 a receita com publicidade nos jornais online supere a dos tradicionais<sup>3</sup>.

Dentre os fatores que justificam essa migração da utilização dos meios de comunicação mais tradicionais para meios de comunicação mais modernos no que se refere à veiculação de anúncios, podem ser destacados: *(i)* o aumento no número de pessoas expostas a esses meios de comunicação; *(ii)* sua maior capacidade para atrair a atenção da audiência e *(iii)* técnicas mais efetivas para mensuração da efetividade dos anúncios [3].

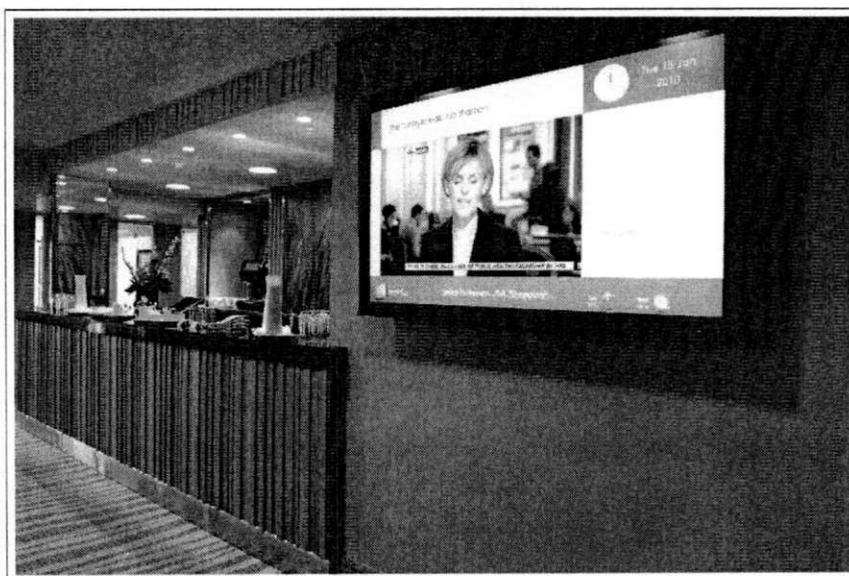


Figura 1.1: Exemplo de sinalização digital.

Dentre esses novos meios de comunicação, a Sinalização Digital tem se destacado Figura 1.1. Ela é considerada uma evolução e adaptação da publicidade tradicional usada em ambientes internos Figura 1.2 e ao ar livre Figura 1.3, como pôsteres e painéis estáticos, às tendências atuais [25]. A sinalização digital difere desses meios por permitir a transmissão

<sup>1</sup><http://www.mediapost.com/publications/article/149737/>

<sup>2</sup><http://mediadecoder.blogs.nytimes.com/2010/11/17/cord-cutting-cable-subscriptions-drop-again/>

<sup>3</sup><http://www.baguete.com.br/noticias/internet/12/04/2011/publicidade-na-web-superara-jornais-em-2013>

dinâmica de conteúdos através de painéis eletrônicos que variam em relação a seu tipo, sendo mais comum o uso de televisores LCD ou de Plasma. Recentemente há uma expansão em seu uso influenciada diretamente pelo barateamento dessas tecnologias [29], potencial publicitário [3] e a diminuição dos custos associados à manutenção e logística quando comparado à mídia indoor, seu concorrente direto [14].

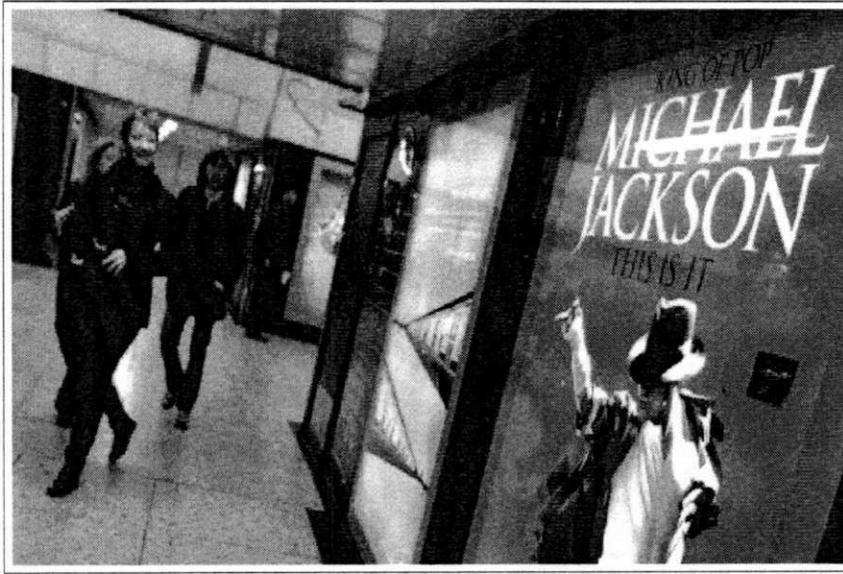


Figura 1.2: Exemplo de publicidade interna.



Figura 1.3: Exemplo de publicidade ao ar livre.

As tecnologias utilizadas pela sinalização digital possibilitam o uso de recursos multimídia como áudio e vídeo que oferecem maiores atrativos para captar a atenção da audiência.

Quando o painel eletrônico está conectado a um computador há a possibilidade de maximizar esses atrativos ao permitir criar conteúdos que interagem com a audiência, usar sensores para analisar as características da audiência, atualizar dinamicamente o conteúdo do painel, dentre outros benefícios [3]. Além disso, os painéis podem ser instalados nas mais diversas situações e locais, como em elevadores, centros comerciais, restaurantes, supermercados, aeroportos, bancos e até locais mais inusitados como banheiros. Essas características fazem com que a sinalização digital esteja presente no cotidiano das pessoas, sendo classificada no conceito de publicidade pervasiva ou *pervasive advertising* [17].

A publicidade pervasiva baseia-se em alguns dos conceitos propostos por Weiser em 1988 para a computação pervasiva [45]. Esse paradigma vislumbra uma sociedade onde os dispositivos tecnológicos estariam presentes no dia a dia das pessoas, nos mais variados objetos, e o seu contato com o usuário seria feito sem influenciar seus hábitos. Para a publicidade esses dispositivos representam uma oportunidade para exibir anúncios em locais antes inacessíveis aos meios de comunicação tradicionais (Figura 1.4). Apenas com a sinalização digital estima-se que sua receita global com a exibição de publicidade teve um crescimento de 16.9% em relação a 2010, alcançando o valor aproximado de \$7,56 bilhões de dólares em 2011.



Figura 1.4: Exemplo de sistema de sinalização digital instalado em um banheiro.

## 1.1 **Problemática**

Apesar de oferecer um grande diferencial para a área de publicidade, a veiculação de anúncios através da sinalização digital possui problemas semelhantes aos dos outros meios de comunicação de massa [31]. Um exemplo é a necessidade de identificação da audiência para exibir anúncios adaptados a seus interesses [36]. Pessoas estão expostas a milhares de anúncios diariamente, porém, de maneira inconsciente, ignoram ou abstraem a maioria deles. Por exemplo, ao acessar uma página Web, os usuários costumam não prestar atenção a áreas usualmente destinadas aos anúncios. Esse processo é conhecido por atenção seletiva [30]. Esse problema se propaga aos anunciantes, pois paga-se por um espaço onde seu produto é exposto a pessoas que não são seu público-alvo. Por outro lado, pesquisas apontam que boa parte da atenção seletiva acontece pela falta de anúncios relevantes, ou seja, quando os anúncios trazem produtos que despertam o interesse dos consumidores sua atenção aumenta [8].

A relevância de um anúncio difere de pessoa para pessoa e não é um processo simples de ser determinado. A combinação de diferentes informações, algumas externas ao consumidor, é que produzirá essa percepção [22]. Por exemplo, um anúncio de um relógio feminino exibido para uma mulher possivelmente terá uma percepção de relevância diferente se fosse exibido para um homem. No contexto de ambientes públicos, como é o caso da sinalização digital, esse processo é ainda mais complexo, pois a relevância do anúncio será baseada no interesse de diferentes pessoas simultaneamente [16]. Pela interação mais casual que há entre esse meio e a audiência, são menores as opções para obter respostas da audiência que possibilitem inferir seus interesses e nem sempre é possível contar com o retorno do usuário, conseqüentemente aumentando a incerteza do assunto [36]. Por exemplo, uma pessoa pode ter contato com um painel instalado em uma determinada loja e por algum motivo não retornar mais, tornando mais difícil a criação de um perfil.

Em resumo, somente através da exibição de anúncios relevantes será possível atrair a atenção dos consumidores. Para isso é importante que os sistemas contem com mecanismos para obter informação dos consumidores. Os trabalhos encontrados na literatura tentam minimizar esse problema solicitando explicitamente dados da audiência ou utilizando dispositivos para identificar o contexto, como sensores, dispositivos móveis ou câmeras. Esses

fatores podem afetar a viabilidade desses projetos, seja economicamente ou pela falta de colaboração da audiência [23] [28]. Diante disso, surge a seguinte questão de pesquisa: *Como aumentar a efetividade na seleção de anúncios em um cenário de sinalização digital sem contar com informações explícitas da audiência?*

Para responder a essa pergunta, acredita-se que o uso de elementos relacionados à teoria do marketing, em específico de estudos na área de comportamento do consumidor e do processo de segmentação de mercado, possam ser usados para determinar a relevância dos anúncios. Espera-se que o uso de conhecimento especialista minimize a necessidade de interação com a audiência e a dependência de dispositivos exclusivamente criados, ou adaptados, para esse fim.

O estudo do processo de compra revela que as aquisições do consumidor fornecem indícios sobre o seu perfil. Esse fato foi evidenciado por um estudo preliminar realizado, que mostra a viabilidade em obter informações sobre o consumidor a partir do estabelecimento onde ele está. Por exemplo, um centro de estética costuma ser frequentado por mulheres. No cenário da sinalização digital, como há uma relação direta com o estabelecimento onde o painel está instalado, propõe-se usar os dados de marketing desse local para inferir o perfil do consumidor. Uma vez identificado o perfil do público-alvo do estabelecimento é possível determinar o que é de interesse mútuo de seus consumidores, conseqüentemente anúncios relacionados a essa característica possuirão maiores chances de serem considerados relevantes [38]. Por exemplo, um painel eletrônico instalado em um salão de exposição de automóveis que faça a exibição de anúncios relacionados a carros possui grandes chances de ser considerado relevante.

Sendo assim, enuncia-se o seguinte problema: *como determinar a relevância do anúncio com base em informações sobre o local onde a sinalização digital está instalada, eliminando a necessidade de interação com a audiência.*

## 1.2 Objetivos

### 1.2.1 Objetivo geral

Neste trabalho, tem-se como objetivo desenvolver um método para melhorar a efetividade da seleção de anúncios em um cenário de sinalização digital. As entradas desse método são informações de marketing relacionadas ao público-alvo e categorias de interesse dos anúncios e estabelecimentos onde os painéis se encontram. Serão selecionados os anúncios mais relacionados aos interesses do estabelecimento ou mais próximos ao seu público-alvo, aumentando as chances de serem considerados relevantes.

O método é fundamentado na combinação dos resultados de dois métodos distintos. O primeiro, seleciona anúncios com base na similaridade existente entre os interesses dos anúncios e do estabelecimento. O segundo utiliza os conceitos da segmentação de mercado para encontrar os anúncios direcionados ao público-alvo do estabelecimento, selecionando o anúncio que for mais próximo.

As categorias de interesse e público-alvo dos produtos e estabelecimentos são a base de qualquer empresa e em geral são definidas antes mesmo da criação de um produto [30]. O fácil acesso a essas informações foi um fator determinante para sua escolha e espera-se que isso reduza as dificuldades de adoção do método proposto.

Os conhecimentos de especialistas em publicidade serão usados para definir os dados de marketing dos anúncios e estabelecimentos. Para ambos será definido o segmento de mercado em que atuam e as categorias de interesses associadas.

A realização de um experimento possibilitará a validação do método proposto, avaliando sua efetividade em comparação a outros métodos, com propósitos similares, encontrados na literatura ou usados em sistemas de sinalização comerciais.

### 1.2.2 Objetivos específicos

Para alcançar o objetivo geral, têm-se como objetivos específicos:

- Identificar variáveis do comportamento do consumidor e processo de compra que podem ser usadas no contexto deste trabalho;

- Definir quais das variáveis identificadas podem ser usadas no contexto do trabalho de modo implícito e sem intervenção tecnológica;
- Definir como essas variáveis serão modeladas computacionalmente e como será o processo de recomendação;
- Planejar a arquitetura para a implementar o método proposto;
- Classificar os anúncios e estabelecimentos em relação a suas categorias de interesse e público-alvo;
- Desenvolver um estudo de caso para validar o trabalho;
- Realizar experimentos para validar o trabalho.

### 1.3 Relevância

Uma boa comunicação com a audiência é quem definirá o sucesso ou insucesso de uma organização [2]. Com o consumidor mais individualista e exigente, além da inúmera quantidade de meios de comunicação, cresce a demanda por conteúdo relevante às suas necessidades em canais de comunicação oportunistas, aproveitando situações onde outros meios de comunicação não atuam.

A sinalização digital é um meio de comunicação capaz de alcançar a audiência em locais antes inacessíveis aos meios de comunicação tradicionais. Por estar apoiada em recursos tecnológicos, como computadores e painéis eletrônicos, há a possibilidade de adaptar dinamicamente seu conteúdo para apoiar as necessidades da audiência. Pela demanda exposta, faz-se necessária a proposta de um método para selecionar anúncios relevantes nesse meio de comunicação.

Dos trabalhos encontrados na literatura [37], [10], [27], [46], as abordagens propostas dependem de dados da audiência que são obtidos explicitamente, com a audiência interagindo com o painel, ou implicitamente, através de sensores, dispositivos móveis ou webcams. Essas limitações podem inviabilizar esses trabalhos devido à necessidade de participação da audiência, que pode não ocorrer, ou pela aquisição desses dispositivos, o que acarreta um aumento nos custos do sistema. O método proposto tem por objetivo encontrar

anúncios relevantes respeitando essas restrições, para isso utilizará dados de marketing que se adequam ao cenário.

O estudo sobre o comportamento do consumidor e processo de compra possibilita a identificação de variáveis que podem ser usadas para selecionar anúncios mais efetivos, minimizando a interação com a audiência e dependências tecnológicas, além dos que são usados em um sistema de sinalização digital tradicional. Em específico à área de sinalização digital, não foi encontrada na literatura nenhum método que use conceitos de marketing na seleção de anúncios. Além disso, a base do método proposto utiliza conceitos que poderão ser adaptados a outros cenários de recomendação.

O experimento realizado suporta as afirmações de marketing sobre a segmentação de mercado e encontrada também em outros trabalhos [11] [34] de que as aquisições do consumidor fornecem indícios sobre suas características e preferências. Essa contribuição é importante para a área de marketing e computação, pois fortalece o uso dessa informação em outros trabalhos.

Finalizando, esta dissertação faz parte do projeto PerComp, do Laboratório de Sistemas Embarcados e Computação Pervasiva da Universidade Federal de Campina Grande. Espera-se que a contribuição da pesquisa auxilie novas pesquisas na área de publicidade pervasiva dentro da instituição.

# Capítulo 2

## Fundamentação teórica

### 2.1 Computação pervasiva

Computação pervasiva ou ubíqua é um termo criado em 1988 por Mark Weiser para designar um novo paradigma da computação [45]. A base deste modelo está no uso de diversos dispositivos tecnológicos simultaneamente, através da sua integração em objetos de nosso dia a dia, de modo que sua presença ou interação seja feita de modo intuitivo ou até mesmo invisível. Apesar de haver uma complexa troca de informações entre os dispositivos e as redes, para o usuário esse processo é omitido, minimizando sua necessidade em aprender determinada tecnologia ou como ela funciona para utilizá-la.

Um ambiente pervasivo pode ser descrito com base nos seguintes sistemas e tecnologias [43]:

- Embarcado: dispositivos integrados no ambiente;
- Ciente de contexto: dispositivos que reconhece o ambiente e dados referente a ele;
- Personalizado: se adaptam as necessidades dos usuários;
- Adaptativos: respondem as mudanças no ambiente;
- Antecipativos: se antecipam as necessidades dos usuários sem necessidade de interação explícita.

Pode-se exemplificar um ambiente pervasivo com o exemplo de uma casa equipada com múltiplos sensores. Nela há um sensor de presença para detectar quando o dono da residência

chega em casa. Essa informação é repassada a uma geladeira que coloca sua cerveja para gelar, pois nela há um sensor que reconhece o padrão de consumo dessa pessoa.

## 2.2 Publicidade pervasiva

A computação pervasiva apresenta diversas possibilidades de aplicações que podem aproveitar as características desse paradigma para inovar em suas áreas. Inovar não é apenas propor algo diferente, mas também se adequar às futuras necessidades das pessoas. Recentemente uma aplicação que vem se destacando é transmissão de anúncios publicitários em dispositivos com tecnologia pervasiva. A esse segmento de aplicação é dado o nome de publicidade pervasiva.

Pelo uso de recursos tecnológicos, com fácil interação e por estar integrada em objetos do cotidiano, a computação pervasiva se apresenta como um meio ideal para a exibição de publicidade [26]. Através dela é possível criar anúncios mais atrativos à audiência e exibí-los em locais antes inacessíveis aos meios de comunicação tradicionais.

Em [35] são descritos alguns desafios à publicidade pervasiva:

- *Entregar o anúncio certo para a pessoa certa*: encontrar os anúncios que mais se adequem as necessidades dos consumidores;
- *Entregar o anúncio no momento correto*: além da restrição acima, o anúncio deve ser exibido no momento correto, por exemplo, um anúncio de pizza ou cerveja teria uma maior relevância se exibido a um grupo de homens assistindo a um jogo de futebol;
- *Entregar o anúncio no meio de comunicação correto*: havendo a interação entre diversos dispositivos, o anúncio deveria ser exibido no dispositivo mais adequado ao momento. Por exemplo, um anúncio mais personalizado aos interesses de uma pessoa poderia ser exibido em seu celular, enquanto para um grupo poderia ser exibido em um painel de sinalização digital;
- *Exibir anúncios inovadores*: inovar na exibição de anúncios, indo além dos interesses tradicionais dos consumidores;

- *Prover meios para a audiência obter mais informações sobre o anúncio:* integrar outros formatos de comunicação para que a audiência possa obter mais informações sobre os anúncios, por exemplo, associar um QR Code ao anúncio para direcioná-lo ao site da organização;
- *Definir formatos de cobrança do espaço publicitário:* definir como o espaço publicitário será vendido;
- *Privacidade e segurança:* a privacidade das informações é um aspecto fundamental aos sistemas pervasivos, por essa razão esse dado não pode ser exposto publicamente.

Pelo uso de recursos tecnológicos inovadores e estar nos mais variados ambientes, esses fatores tornam a publicidade pervasiva um meio de comunicação único. A citar algumas de suas vantagens [26]:

- *Comunicação simétrica:* na publicidade tradicional os anunciantes definem o que será visto e quando será visto, limitando a capacidade de escolha dos consumidores. Pela característica interativa da computação pervasiva, a audiência possui meios para comunicar sua opinião e influenciar nos anúncios que serão escolhidos;
- *Simplificação do modelo de anúncios:* pelo uso de dispositivos tecnológicos a publicidade pervasiva reduz os custos para a produção de anúncios, facilitando o acesso de pequenas empresas ao mercado publicitário;
- *Novas experiências:* anúncios com gráficos em alta definição e áudio são exemplos de recursos acessíveis à publicidade pervasiva e que tornam esse meio muito mais atrativo à audiência;
- *Personalização e adaptação ao contexto:* personalização é o núcleo da computação pervasiva. Através dela é possível exibir os anúncios mais adequados aos interesses dos consumidores;
- *Mensuração da audiência:* para a publicidade medir a receptividade de um anúncio é fundamental para tomada de decisões sobre como melhorá-lo. Na publicidade tradicional o uso de sensores capazes de avaliar quem são as pessoas expostas aos anúncios

e como elas reagiram a eles é mais complexo e em alguns casos inviável. Para a publicidade pervasiva esse é um procedimento simples:

- *Persuasão automática*: modificar as atitudes da audiência faz parte da publicidade. O uso de computadores facilita esse processo, possibilitando conhecer melhor quem são os consumidores e sugerir conteúdo que possa alcançar esse objetivo.

Nas pesquisas recentes em publicidade pervasiva destacam-se o uso de dispositivos móveis, sinalização digital e dispositivos físicos de computação, como QR Codes ou NFC [26]. Esses dispositivos não possuem todas as características ubíquas, mas atualmente são as alternativas mais próximas. Quando combinados, consegue-se uma maior interação e engajamento da audiência próximo a um ambiente pervasivo.

## 2.3 Sinalização digital

A sinalização digital é um meio de comunicação que utiliza painéis eletrônicos, em geral televisores com tecnologia LCD ou Plasma, para transmitir sua mensagem. Classificada no segmento de mídias *Out-Of-Home* (OOH), a mensagem da sinalização digital é passada para a audiência enquanto elas estão fora de suas residências, em espaços públicos, como no trânsito, a espera de um atendimento, ou mesmo em locais mais inusitados como no ônibus ou elevadores (Figura 2.1). Essa integração no cotidiano das pessoas, associado ao uso de recursos tecnológicos, faz da sinalização digital um dos principais meios usados pela publicidade pervasiva [26].

Há três classificações para a sinalização digital, elas variam em razão do local onde o painel é instalado e da forma como ele é utilizado [21]:

1. *Pontos de interesses (Point Of Interest)*: nesse cenário os painéis eletrônicos estão instalados próximos a locais onde os consumidores tomam suas decisões. Seu objetivo é reforçar ou modificar as intenções do consumidor em relação a algum produto.
2. *Ponto de passagem (Point Of Transit)*: painéis são instalados em locais de grande tráfego e seu objetivo é atrair a atenção da audiência por um breve período de tempo. Como exemplo, um painel instalado em uma rodovia.



Figura 2.1: Sinalização digital instalada em ônibus [28].

3. *Ponto de espera (Point Of Wait)*: direcionada à audiência que aguarda por algum produto ou serviço. A motivação para o seu uso é que essas pessoas, por estarem ociosas, estão mais receptivas a mensagens. Esse tipo de painel está localizado normalmente em filas de espera em lojas, hospitais e filas de banco.

A sinalização digital é considerada uma evolução da mídia OOH tradicional, como pôsteres e *outdoors*. Nesses meios de comunicação a troca de conteúdo é um processo lento e caro em termos de mão de obra e materiais utilizados [14]. Há custos envolvidos com a fabricação do material no qual o conteúdo é impresso; posteriormente, sua distribuição e instalação também demandam gastos e requerem uma logística para a troca dos conteúdos previamente instalados por novos. O processo é denominado ciclo de criação, distribuição e instalação do anúncio (Ciclo CDI), ilustrado na Figura 2.2.

Por usar uma tecnologia que permite a troca rápida de conteúdo, a sinalização digital provê redução dos custos associados a esse processo e se mostra muito viável para a publicidade. Com exceção do processo de criação, onde é desenvolvida a arte gráfica do anúncio, a sinalização digital simplifica os processos de distribuição e instalação de anúncios, haja vista que esses processos podem ser feitos por meio de uma simples troca da mídia que contém os anúncios, por exemplo, um CD ou DVD. Com a possibilidade de se conectar o painel eletrônico à Internet, esses processos podem ser parcialmente eliminados. Nesse caso, a dis-



Figura 2.2: Processo de criação, distribuição e instalação de anúncios publicitários (Ciclo CDI) [14].

tribuição e instalação dos anúncios são feitas diretamente do servidor de anúncios para os painéis eletrônicos.

Essa facilidade na troca do conteúdo exibido no painel oferece a sinalização digital diversas vantagens em comparação aos meios tradicionais, dentre elas:

- *Exibição de múltiplos anúncios:* ao contrário dos pôsteres e outdoors que exibem apenas um anúncio por vez, a sinalização digital possibilita que múltiplos anúncios sejam exibidos em um mesmo painel, consequentemente aumentando as receitas com publicidade que esse meio pode gerar;
- *Adaptação do conteúdo:* com o uso de sensores o conteúdo do painel pode ser adaptado às pessoas que estejam presentes em um local ou a variáveis de contexto, como o clima e o horário. Por exemplo, um local que é frequentado no período da manhã por famílias e no período da noite por adultos requer a exibição de diferentes anúncios para cada

grupo;

- *Manter o conteúdo sempre atualizado*: um problema comum ao uso de pôsteres e outdoors é que para anúncios com um prazo de validade, como eventos, é preciso que eles sejam instalados antes do evento ocorrer. Pelos problemas citados anteriormente, podem ocorrer imprevistos no ciclo CDI que inviabilizem colocar os anúncios a tempo. Além disso, após o evento ocorrer o anúncio perde seu sentido. Com a sinalização digital minimiza-se esse problema com a facilidade na atualização do conteúdo exibido.

Além dessas vantagens, para o anunciante a sinalização digital oferece recursos mais efetivos para captar a atenção da audiência, como a possibilidade de usar recursos multimídia de alta definição nos anúncios e meios de interagir com a audiência.

## 2.4 Anúncios publicitários

Para uma empresa não basta apenas comercializar seus produtos ou serviços, é preciso que haja um processo de comunicação com os potenciais consumidores para informa-los sobre os benefícios oferecidos ou tornar os produtos parte de seus interesses. Esse processo de comunicação é feito através dos anúncios publicitários, promoções de venda e relações públicas.

Os anúncios publicitários são uma forma de comunicação, paga, cujo objetivo é persuadir, comunicar ou informar a potenciais consumidores sobre os produtos de uma empresa [30].

O objetivo principal dos anúncios é estimular uma resposta do público-alvo em relação ao que está sendo anunciado. Essa resposta pode ser a aquisição do produto, associação da marca a pessoas, locais, eventos, experiências ou objetos, ou sua lembrança posteriormente, no momento de compra.

Para alcançar esse objetivo a mensagem deve estar relacionada aos interesses da audiência que irá assisti-lo. Esse princípio torna a mensagem relevante a essas pessoas, facilitando sua interpretação e aceitação.

### 2.4.1 Relevancia dos anúncios publicitários

Para a publicidade um anúncio é relevante quando é útil ou sobressai às expectativas dos consumidores. Há teorias que buscam explicar como ocorre esse processo, dentre elas a que mais se destaca é a teoria da relevância proposta por Sperber e Wilson [40].

Essa teoria se baseia em uma característica básica da cognição humana: assuntos de interesse recebem mais atenção, pois requerem um menor esforço no processamento da comunicação. Essa mesma característica é denominada por Kotler como atenção seletiva [30].

Para Sperber e Wilson, do ponto de vista comunicativo, o estímulo usado na mensagem publicitária deve ser adaptado às expectativas do consumidor para atrair sua atenção. Uma alternativa para alcançar esse nível de personalização é identificar as características e necessidades dos consumidores e exibir anúncios relacionados a esses assuntos. Fornecendo essas garantias a mensagem justificará o esforço cognitivo de processamento.

Além do citado, Kotler também descreve alguns fatores que podem influenciar os consumidores a considerar o anúncio relevante:

1. A mensagem deve ser de interesse dos consumidores;
2. A mensagem precisa ser transmitida de modo que crie impacto, utilizando inovações tecnológicas e artísticas, criando um diferencial;
3. A audiência dá mais atenção a anúncios que tragam novas informações;
4. A mensagem deve justificar ou reforçar as recentes decisões de compra do consumidor.  
Por exemplo, uma pessoa que comprou um automóvel provavelmente passará a dar maior atenção a anúncios relacionados a carros. Esse processo é chamado de redução da dissonância cognitiva.

Percebe-se que para exibir anúncios adequados é preciso conhecer quem serão expostos a eles. Esse processo envolve identificar os hábitos desses consumidores e influências que afetam suas escolhas, para que o anúncio possa se adequar ou tirar proveito desses elementos.

Há uma área no marketing destinada unicamente ao estudo desses fatores, é a área de comportamento do consumidor. Ela é responsável por identificar esses fatores e compreender como eles se relacionam e de que forma influenciam o consumidor.

### 2.4.2 Teoria sobre o comportamento do consumidor

Para um anúncio ser relevante é necessário que o consumidor tenha interesse pelo que está sendo anunciado. Para isso é preciso conhecer quem são os atuais e potenciais consumidores do produto anunciado, quando eles compram, onde e como tomam essas decisões.

Conhecer quem é a audiência possibilita compreender quais são seus hábitos, quais valores essas pessoas prezam, quais canais de comunicação costumam usar, dentre outros fatores. Com essas informações as empresas podem adotar ações de marketing para minimizar o impacto desses elementos, caso seja negativo, ou aproveitá-los criando anúncios com maior potencial para serem considerados relevantes e os exibi-los nos canais de comunicação adequados [31].

Os fatores que mais afetam os consumidores são classificados em quatro categorias: elementos culturais, sociais, pessoais e características psicológicas [22].

Os elementos culturais são o fator com influência mais ampla e profunda sobre o comportamento de uma pessoa e são a causa mais básica sobre seu querer e comportamento [31]. Eles se referem à soma complexa de princípios, valores, crenças e tradições adquiridas por uma pessoa com a sociedade, família e outras instituições presentes em sua vida. Por exemplo, parte das roupas utilizadas e hábitos alimentares dos brasileiros diferem dos chineses, indianos ou suecos, pela diversidade cultural que há entre esses países.

Dentro de uma cultura pode haver grupos menores, mais homogêneos, em relação a seus princípios e tradições. Esses grupos são classificados pelo termo de subcultura e muitas vezes apresentam oportunidades importantes de segmentação para o marketing [31]. Enquadra-se dentro da subcultura pessoas de uma mesma região geográfica, religião, grupos raciais e étnicos, dentre outros. Por exemplo, há uma expansão recente no número de produtos direcionados a membros da igreja evangélica pelo recente aumento no número de pessoas filiadas a essas igrejas.

As características pessoais do consumidor também o influenciam em suas escolhas. Fazem parte dessas características seus relacionamentos com amigos e famílias, sua idade, condições financeiras, estilo de vida, personalidade e autoconceito.

Amigos, família e vizinhos exercem influência, porém em diferentes níveis [31]. Um consumidor com esposa e filhos faz suas escolhas ponderando a existência de sua família. Essas escolhas também variam com a idade do consumidor. Elas diferem bastante entre um

consumidor de 10 anos para outro com 30 anos. As escolhas também são limitadas pela condição financeira do consumidor, pois ela será um fator restritivo às opções de compra.

Uma classe de elementos complexos em sua identificação, mas com forte influência nas decisões são os fatores psicológicos. Faz parte dessa categoria a motivação que leva uma pessoa a desejar determinado produto, processo estudado pelas teorias da motivação de Freud e das necessidades de Maslow. Há também aspectos relacionados à percepção do consumidor, que envolve sua capacidade em registrar um anúncio visto ou abstraí-lo. Também fazem parte dessa classe as crenças e atitudes do consumidor. Esses elementos são os mais complexos de identificar, pois variam de indivíduo para indivíduo.

Os fatores apresentados possibilitam aos gestores de marketing identificar características comuns aos consumidores e agrupá-los, criando o conceito de segmentos de mercado. Por exemplo, os consumidores residentes na Índia ou no Brasil, fazem parte de segmentos geográficos distintos. Pode-se agrupá-los também em razão da sua faixa etária, criando dois segmentos para agrupar os consumidores com 20 anos e outro para consumidores de 60 anos.

Esses segmentos ilustram o objetivo da segmentação de mercado, que é o de dividir um mercado heterogêneo em partes mais homogêneas permitindo que ações de marketing possam ser adequadas às necessidades de cada segmento. Quando uma organização define o(s) segmento(s) em que irá atuar, aos membros desse segmento é dado o nome de público-alvo.

### **2.4.3 Segmentação de mercado e público-alvo**

Em cada segmento há pessoas que compartilham características e conseqüentemente possuem necessidades e comportamentos próximos [31]. Através da segmentação de mercado as empresas podem adequar seus produtos às necessidades de cada mercado, ajustando seu preço, design, canais de distribuição e publicidade.

Com base nas variáveis identificadas no estudo do comportamento do consumidor, os gestores de marketing as utilizam, individualmente ou em conjunto, para definir como os segmentos serão formados. As principais variáveis usadas nesse processo são as variáveis geográficas, demográficas, psicográficas e comportamentais. Kotler, em seu livro [31], apresenta um modelo dos possíveis valores para essas variáveis, ilustrado na Tabela 2.1.

Os segmentos formados pelas variáveis geográficas dividem os mercados em unidades

DEMOGRÁFICOS	
Idade	0-6, 6-11, 12-19, 20-34, 35-49, 50-64, 65+
Sexo	Masculino, Feminino
Família	1 ou 2 pessoas, 3 ou 4 pessoas, 5 pessoas ou mais
Renda	0-\$10,000; \$10,000-\$20,000; \$20,000-\$30,000; \$30,000-\$50,000; \$50,000+
Ocupação	Profissional, técnico, gerente, oficial, vendedor, desempregado, etc.
Educação	Analfabeto, Ensino Fundamental, Ensino Médio, Ensino Superior, Pós-Graduado
Religião	Católico, Protestante, Judeu, Muçulmano, Hindu, entre outras
Raça	Asiáticos, Hispânicos, Negros, Brancos
Nacionalidade	Norte-Americanos, Sul-Americanos, Britânicos, Franceses, Italiano, Japoneses
PSICOGRÁFICOS	
Classe Social	A1, A2, B1, B2, C1, C2, D, E
Estilo de vida	Realizadores, Esforçados, Sobreviventes
Personalidade	Compulsiva, Autoritária, Ambiciosa
GEOGRÁFICOS	
Região Mundial	América do Norte, Oeste Europeu, Ásia
País	Canadá, Alemanha, China, México, Brasil
Densidade	Urbano, Suburbano, Rural
Clima	Tropical, Subtropical, Equatorial, Mediterrâneo, etc.
COMPORTAMENTAL	
Ocasões	Ocasão Regular; Ocasão Especial
Benefícios	Qualidade, Serviço, Economia, Conveniência, Velocidade
Status do usuário	não-usuário, ex-usuário, usuário em potencial, primeira vez, usuário regular
Lealdade	Nenhuma, Média, Forte, Absoluta
Conhecimento	Não Conhece o Produto, Consciente, Informado, Interessado, Desejoso
Atitude em relação ao produto	Entusiasta, Positiva, Indiferente, Negativo, Hostil

Tabela 2.1: Variáveis usadas no processo de segmentação de mercado.

geográficas, como países, estados, cidades e vizinhanças. Apesar de haverem produtos que se adequam a muitos mercados, há especificidades para esses consumidores que justificam alguma forma de personalização. Essa é uma prática muito adotada por fabricantes de automóveis que dividem seus mercados pelas regiões geográficas dos seus consumidores. Os carros vendidos nos Estados Unidos são diferentes dos carros vendidos no Brasil ou no Japão. As demandas por segurança, design e estilo de carro variam entre esses países, justificando a necessidade de adaptação. Diferenças climáticas também entram nessa categoria, fazendo com que alguns produtos sejam mais adequados a climas frios enquanto outros a climas mais quentes.

A segmentação demográfica divide o mercado com base em variáveis como faixa etária, gênero, orientação sexual, tamanho da família, renda, religião, dentre outros. Esse formato de segmentação é o mais usado pelas empresas, principalmente por duas razões: *i*) as necessidades e desejos dos consumidores variam diretamente pelas mudanças nessas variáveis, e *ii*) é uma informação fácil de obter e ser aplicada [31]. Quando essas variáveis são combi-

nadas têm-se segmentos com características e necessidades bem definidas. Por exemplo, há variações distintas nos desejos de um segmento formado por homens de 30 a 40 anos quando comparado a um segmento de mulheres com a mesma faixa etária. Cada segmento justifica diferentes anúncios, por exemplo, exibir anúncios de carros esportivos para os homens possivelmente seria mais eficiente que exibir um anúncio de tratamento capilar, enquanto para as mulheres possivelmente seria a situação inversa.

Outra forma eficaz para agrupar consumidores com necessidades semelhantes é o uso de variáveis como a classe social, estilo de vida e características pessoais. Essas variáveis influenciam diretamente nas escolhas do consumidor, que são usadas para sustentar um status na sociedade, por exemplo, ou para sustentar determinado estilo de vida. Como exemplo uma pessoa que mantém hábitos saudáveis de alimentação e pratica esporte está mais inclinada a se interessar por anúncios de academia ou suplementos alimentares. Os mercados divididos por essas variáveis recebem o nome de segmentos psicográficos.

Por fim, a última categoria de segmentos é denominada segmentação comportamental e agrupa os consumidores com base em seus hábitos, por exemplo: nas ocasiões de como os produtos são usados, nos benefícios pretendidos com ele, em sua frequência de uso, lealdade, preparação para comprá-lo e atitudes em relação ao produto.

Através da segmentação de mercado é possível identificar quem são os potenciais consumidores de um produto, seus hábitos, necessidades e canais de comunicação que utilizam. Essa informação pode ser usada para personalizar os anúncios que serão exibidos, escolhendo aqueles que tenham alguma relação com o público-alvo aos quais essas pessoas pertencem.

## **2.5 Lógica nebulosa**

A lógica nebulosa surge como uma alternativa à lógica booleana para possibilitar o tratamento das informações quando elas são incertas ou inexatas. Ao contrário da teoria clássica dos conjuntos onde há apenas a possibilidade de um elemento pertencer ou não pertencer a um conjunto, a lógica nebulosa é mais flexível e possibilita que um elemento apresente diferentes níveis de pertinência aos conjuntos. Ela segue os princípios da teoria dos conjuntos nebulosos, proposta em 1965 por Lotfi Zadeh [20]. A flexibilidade da teoria dos conjun-

tos nebulosos associada ao raciocínio nebuloso permite criar sistemas que se comportem de modo semelhante ao ser humano quando este raciocina com base em fatos vagos.

Neste estudo a incerteza está presente quando é preciso classificar uma pessoa como pertencente ao público-alvo de um anúncio. Caso todos os dados dessa pessoa satisfaçam as restrições desse público-alvo, ela é classificada como pertencente a ele. Por exemplo, um anúncio de perfume direcionado a mulheres de 15 a 40 anos. Um consumidor do sexo feminino com 20 anos faz parte desse público-alvo. No entanto, em situações onde apenas parte dos dados satisfazem as restrições, faz-se necessário definir níveis intermediários que possibilitem representar essas classificações.

Nessa dissertação a lógica nebulosa será usada para classificar a proximidade entre o público-alvo dos anúncios que estarão disponíveis para exibição e o público-alvo do estabelecimento onde o painel está instalado. Como o público-alvo desse local descreve as características dos potenciais consumidores dos produtos oferecidos no estabelecimento, são grandes as chances das pessoas que estão presentes nele terem essas características.

No processo de classificação, quanto maior a proximidade entre os públicos-alvo do anúncio e do estabelecimento maior a possibilidade do anúncio ser relevante para os consumidores desse local. Por exemplo, o anúncio do automóvel Land Rover (Figura 2.3) é direcionado para homens, entre 34 a 54 anos e com um alto poder financeiro<sup>1</sup>. Pela lógica nebulosa é possível modelar o grau de pertinência de um estabelecimento com público-alvo de pessoas de 20 a 25 anos, idade que não faz parte do público-alvo da Land Rover, mas que possui poder econômico para comprá-lo.

### 2.5.1 Propriedades

Alguns elementos da teoria clássica também se aplicam, com modificações, à teoria dos conjuntos nebulosos, dentre eles: a noção de conjuntos, subconjuntos e elementos. Outros conceitos são introduzidos, como: o grau de pertinência, função de pertinência, variáveis e termos linguísticos, processos de fuzificação e defuzificação, e o raciocínio nebuloso.

<sup>1</sup>[http://www.admob.com/marketing/pdf/LandRover\\_AdMobCaseStudy.pdf](http://www.admob.com/marketing/pdf/LandRover_AdMobCaseStudy.pdf)

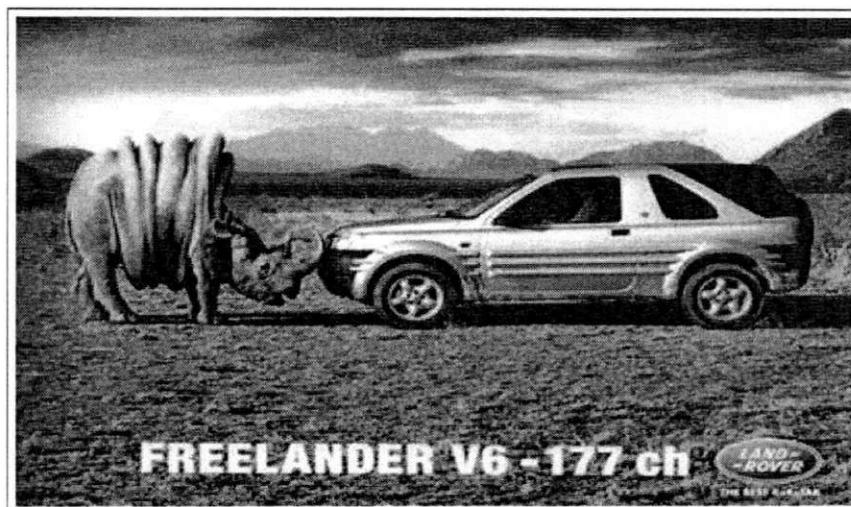


Figura 2.3: Anúncio do veículo Land Rover Freelander.

### Conjunto nebuloso e grau de pertinência

Na lógica clássica os conjuntos possuem limites precisos e um elemento é dito pertencente ou não pertencente a ele. Para a lógica nebulosa essa definição é mais flexível, havendo a possibilidade de um elemento pertencer parcialmente a um ou mais conjuntos. A essa escala de associação é dado o nome de grau de pertinência.

O grau de pertinência de um elemento a um conjunto varia em uma escala de 0 a 1, o valor zero significa que ele não pertence e o valor um significa que ele pertence por completo ao conjunto. Quando um elemento está em um desses extremos é um caso de lógica clássica, nas demais situações esse valor representa uma pertinência parcial ao conjunto. Um diagrama de Venn para um conjunto nebuloso é representado na Figura 2.4. A parte sombreada do círculo é usada para representar os níveis de pertinência de um elemento.

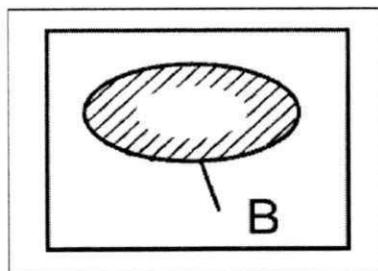


Figura 2.4: Diagrama de Venn para um conjunto nebuloso.

### Função de pertinência

Para calcular o grau de pertinência utiliza-se uma fórmula denominada função de pertinência. Essa função mapeia os elementos ao conjunto em uma escala de 0 a 1. Sendo  $X$  o universo de discurso,  $x$  um elemento desse universo e  $A$  seja um conjunto nebuloso, a função de pertinência que caracteriza esse conjunto é definida como:

$$\mu_A(x) : X \rightarrow [0, 1]$$

As funções de pertinência podem assumir diferentes formatos a depender do contexto em que são usadas. As formas mais comuns são: linear, triangular, trapezoidal, gaussiana e sigmoide.

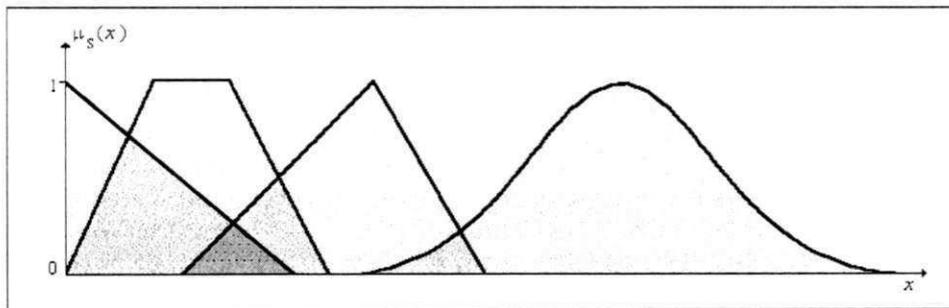


Figura 2.5: Funções de pertinência. [4].

### Variáveis e termos linguísticos

Uma das principais motivações para o uso da lógica nebulosa é a possibilidade de modelar a incerteza associada à vagueza das informações. Há dois conceitos fundamentais a teoria dos conjuntos nebulosos que são usados para descrever essa incerteza, são as variáveis e termos linguísticos associados ao domínio de um problema.

Considere o exemplo anterior da *Land Rover*. Podem-se definir duas variáveis linguísticas que é a idade e o rendimento dos consumidores. Tendo como exemplo um potencial consumidor de 20 anos, a proximidade desse valor com a variável idade que descreve o público-alvo do produto pode ser representada pelos termos linguísticos pertencente, próxima ou excluída. O uso do termo excluído possibilita impor restrições a valores que não

são aceitos. Por exemplo, para um anúncio de cervejas, pessoas abaixo dos 18 anos podem ser excluídas.

Os termos linguísticos representam os conjuntos do sistema nebuloso. Para cada termo há uma função de pertinência para mapear o quanto um valor pertence ao conjunto. O processo que mapeia os dados de entrada em valores de pertinência aos termos linguísticos é denominado fuzificação.

### **Fuzificação e defuzificação dos valores**

A fuzificação é um processo que mapeia os valores de entrada do sistema, conhecido como valores *crisp*, em graus de pertinência em relação aos conjuntos nebulosos. Ou seja, para cada dado de entrada ele é mapeado, através das funções de pertinência, em um valor entre 0 e 1 que demonstra o quanto ele pertence ao conjunto nebuloso.

### **Raciocínio nebuloso**

Uma das maiores motivações para o uso da lógica nebulosa é a sua possibilidade de tirar conclusões ou deduções em um contexto de vagueza. Esse processo, conhecido por raciocínio nebuloso, em sua forma simplista utiliza proposições com um antecedente e uma conclusão, na forma de regras SE e ENTÃO.

Há dois tipos de proposições: simples e compostas. As proposições simples são no formato  $x \text{ é } A$ , onde  $x$  é uma variável linguística e  $A$  é um termo linguístico. Enquanto que as proposições compostas são formadas por proposições simples agrupadas através de operadores E, OU e NÃO. Essas regras nebulosas são conhecidas por regras de *Zadeh-Mamdani* [20].

Como descrito, para descrever o público-alvo dos anúncios e estabelecimentos serão usadas as variáveis idade, gênero e rendimento. Os dados do público-alvo dos anúncios serão comparados aos dados do público-alvo do estabelecimento e essa classificação resultará na possível relevância que o anúncio terá para as pessoas presentes nesse local. Quanto mais próximo os valores do público alvo de um anúncio sejam do público-alvo do estabelecimento, será considerado uma maior possibilidade do anúncio ser considerado relevante.

Supondo que há um painel eletrônico instalado em uma loja da grife de óculos Ray Ban, seu público-alvo é formado por pessoas de ambos os generos, com idade entre 18 a 34

anos e de classe econômica A2, A1 e B1 <sup>2</sup>. Como definir se um anúncio da grife *Tommy Hilfiger*, com um público-alvo de pessoas de faixa etária de 18 a 25 anos e rendimento A2 e A1, seria relevante? Para essa inferência propõe-se utilizar regras de inferência nebulosa, onde os valores que descrevem o público-alvo do anúncio serão classificados em relação às variáveis nebulosas que descrevem o público-alvo do estabelecimento. Quanto mais próxima for essa classificação, maior a possibilidade do anúncio ser considerado relevante. Para cada estabelecimento especialistas definirão as regras de inferência. Cada regra é específica para o contexto do estabelecimento, pois em algumas situações uma variável pode exercer maior influência que outra. Por exemplo, na compra de uma Ferrari a classe social possui mais importância que a faixa etária. Abaixo um exemplo de regra:

*SE idade é próxima E rendimento é pertencente ENTÃO relevância é ALTA*

A inferência nebulosa considera todas as regras especificadas e o resultado final é formado pela agregação dos resultados de cada regra. Por fim, o valor obtido é um número nebuloso que precisa ser transformado em um número crisp, em um processo denominado defusificação. Esse número crisp será a relevância do anúncio para o público-alvo do estabelecimento. O método mais comum para realizar a defusificação é utilizar do centro de gravidade do conjunto nebuloso obtido pelo processo de agregação dos resultados das regras de inferência nebulosa [20].

## **Coeficiente de similaridade de Jaccard**

O índice de similaridade de Jaccard é um cálculo estatístico que possibilita medir a similaridade entre conjuntos de informações assimétricas ou binárias. O conceito de similaridade é fundamental para áreas como mineração de dados, inteligência artificial e biologia, pois possibilita a medição de padrões entre objetos.

Há duas formas para calcular o índice de Jaccard. Quando os valores não são binários pode-se calcular a similaridade para dois conjuntos A e B através da intersecção dos conjuntos dividida por sua união:

<sup>2</sup><http://www.slideshare.net/kaylainanc/ad-plan>

$$\text{sim}(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Quando os elementos dos conjuntos são binários utiliza-se a seguinte função:

$$\text{sim}(A, B) = \frac{p}{p + q + r}$$

Onde:

- $p$  é igual ao total de variáveis que tem o mesmo valor para ambos os conjuntos;
- $q$  é igual ao total de variáveis que tem o valor 1 para o conjunto A e 0 para o conjunto B;
- $r$  é igual ao total de variáveis que tem o valor 0 para o conjunto A e 1 para o conjunto B.

Ambas retornam um número entre 0 e 1 que mede o nível de similaridade entre os conjuntos. O valor de 1 corresponde aos dois conjuntos serem semelhantes enquanto que 0 indica uma dissimilaridade completa entre eles. No contexto desse trabalho, o índice de Jaccard será utilizado para medir a similaridade entre as categorias de interesse associadas ao estabelecimento onde o painel eletrônico está instalado e as categorias de interesse associada aos anúncios a serem exibidos. Esse processo é uma alternativa ao cálculo da relevância pelo público-alvo, possibilitando usar uma informação que descreve bem as necessidades dos consumidores [30].

## Capítulo 3

### Trabalhos relacionados

#### 3.1 MobiDIC

Em [28] é proposto um sistema que utiliza informações de contexto, como dados geográficos e o horário, para melhorar a seleção de anúncios em sinalização digital. O algoritmo identifica automaticamente o melhor anúncio para uma determinada localidade geográfica e horário a partir da resposta da audiência. Para isso são exibidos anúncios com códigos promocionais. Quando os consumidores usam esses códigos o algoritmo tem a informação sobre onde o anúncio foi exibido e seu horário. Esse processo se repete até que o algoritmo tenha dados suficientes para iniciar suas recomendações. Para validar o algoritmo desenvolveu-se um estudo de caso em um sistema de sinalização digital instalado na cidade de Munique.

O algoritmo depende de entradas que são obtidas quando a audiência requisita os descontos promocionais. Associado ao anúncio está um código que identifica o local onde o anúncio foi visto e o horário em que isso ocorreu. Para requisitar esses descontos a pessoa deve fotografar o anúncio e apresentá-lo no estabelecimento. No estabelecimento esse cupom é cadastrado, esse processo serve para associar as variáveis local e horário ao anúncio. A partir dessas informações o algoritmo irá prever qual anúncio exibir diante de um determinado contexto.

O algoritmo utiliza um modelo de venda de espaço publicitário nos painéis eletrônicos, semelhante ao pay-per-click adotado por vários mecanismos de espaço publicitário, dentre eles o Google AdSense. Para uma dada ação  $j$  do usuário, como por exemplo, a conversão do cupom, o anunciante define o quanto ela lhe é interessante  $uj$ . Dado um determinado

conjunto de variáveis de contexto  $F_1...F_n$  e a informação sobre seus valores  $f_1...f_n$ , é predita a quantidade de pessoas que fariam a ação  $j$ , se o conteúdo fosse exibido. O cálculo dessa predição utiliza o histórico de contextos, inseridos no sistema cada vez que um cupom promocional é convertido. Esse cálculo é feito através de um classificador Bayesiano.

Para validar o estudo, um experimento foi conduzido durante aproximadamente um ano. Foram usados vinte painéis eletrônicos instalados nos telefones públicos em diferentes pontos da cidade de Munique (Figura 3.1). Dezessete estabelecimentos participaram do experimento fornecendo anúncios promocionais.



Figura 3.1: Sinalização digital utilizada no estudo e exemplo do cupom exibido [28].

Apesar do longo tempo de execução do experimento, seus resultados não foram satisfatórios, pois apenas 37 cupons de promoção foram resgatados. Essa quantidade de dados foi insuficiente para o funcionamento do algoritmo de seleção do conteúdo. Para entender o porquê da baixa adesão foram realizadas entrevistas com os lojistas e com as pessoas que usaram o sistema. Chegou-se à conclusão de que o baixo número de participantes pode ser explicado pela falta de divulgação do sistema e falha no formato de inserção dos códigos no sistema, pois em alguns estabelecimentos os lojistas não realizaram esse processo, que era manual.

Apesar da incapacidade do algoritmo em gerar recomendações, o trabalho trouxe boas contribuições. A partir das entrevistas foram identificados elementos que influenciam as

expectativas dos consumidores e dos anunciantes em relação à aceitabilidade da proposta. Além disso, foram obtidas sugestões para melhoria do sistema que podem ser aplicadas a outros sistemas de sinalização digital.

O método é um dos poucos encontrados na literatura que conseguiu ser testado em um sistema de sinalização digital em uso. O fato de o experimento ter tido um longo período de duração é mais um ponto a seu favor.

Diferente do método proposto nesta dissertação, o algoritmo de seleção do MobiDic necessita de dados explícitos da audiência para funcionar. Os resultados do experimento são um indício de que em algumas ocasiões essa pode não ser uma alternativa viável. Além disso, há também a necessidade da audiência utilizar algum dispositivo tecnológico para fornecer a resposta necessária ao funcionamento do sistema.

## 3.2 **BluScreen**

Um dos primeiros trabalhos encontrados na literatura que abordam a melhoria na seleção de anúncios em um cenário de sinalização digital é o *Bluscreen* [39]. *BluScreen* é um *framework* distribuído que tem como objetivo maximizar a exposição da audiência a anúncios ainda não vistos. Para o autor, o fato de não exibir anúncios repetidos torna o painel eletrônico mais atrativo às pessoas.

O sistema identifica a audiência por meio de sensores *Bluetooth* instalados nos painéis. Para cada pessoa identificada é mantido um histórico dos anúncios que ela visualizou. Considera-se que sua presença próxima ao painel indica que o anúncio foi visto. Para gerenciamento dos painéis é usado um mecanismo de leilão para venda de espaço publicitário e uma abordagem multiagentes que representam os anunciantes. Quando o sistema estiver no processo de seleção do anúncio, primeiramente é detectada a audiência próxima ao painel, quanto menos dessas pessoas tenham visto o anúncio, maior será o lance que seu agente fará pelo espaço publicitário (Figura 3.2).

O *BluScreen* tem como motivação não necessitar de informações explícitas da audiência. Argumenta-se que essa abordagem possui como vantagens: *i*) não haver necessidade de informações da audiência; *ii*) não requerer nenhuma ação por parte da audiência e *iii*) não haver necessidade de nenhum software extra.

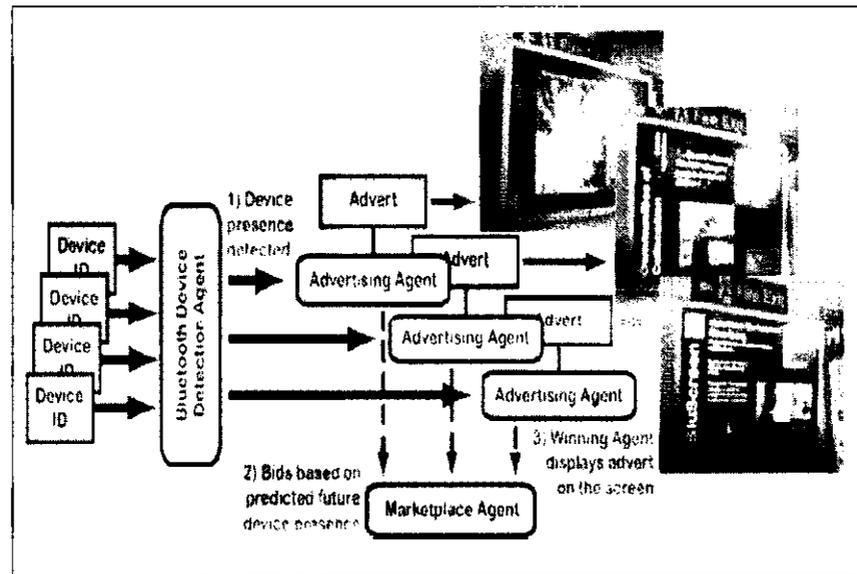


Figura 3.2: Arquitetura da abordagem BluScreen [19].

Dois experimentos foram conduzidos para validar a abordagem. No primeiro validou-se o algoritmo em relação à quantidade de anúncios que seriam necessários exibir para expor a audiência a todos os anúncios cadastrados. Seus resultados foram comparados a outros algoritmos usados na área de sinalização digital. No segundo experimento, o sistema foi avaliado do ponto de vista da sua utilidade para a audiência, considerando o quanto o conteúdo foi interessante para ela.

O primeiro experimento é detalhado em [39] e teve a duração de três meses. Ele possuía dois objetivos: *i*) coletar dados da movimentação de pessoas por uma determinada área, para criar um modelo matemático capaz de prever os caminhos que essas pessoas iriam seguir em um dado momento e *ii*) usar esse modelo para simular a movimentação de pessoas e calcular a probabilidade de, em um dado momento, uma pessoa estar exposta a um painel. Com essa informação, os algoritmos são validados verificando a quantidade de anúncios que precisariam ser exibidos para que todas as pessoas da simulação tenham visto todos os anúncios cadastrados. Compararam-se os resultados do algoritmo proposto aos resultados dos algoritmos de seleção aleatória e seleção sequencial. Os resultados mostram que o BluScreen superou os outros algoritmos ao necessitar de menos exibições para que as pessoas visualizassem todos os anúncios cadastrados (Figura 3.3).

O segundo experimento discutido em [19] teve como objetivo avaliar a utilidade do *BluS-*

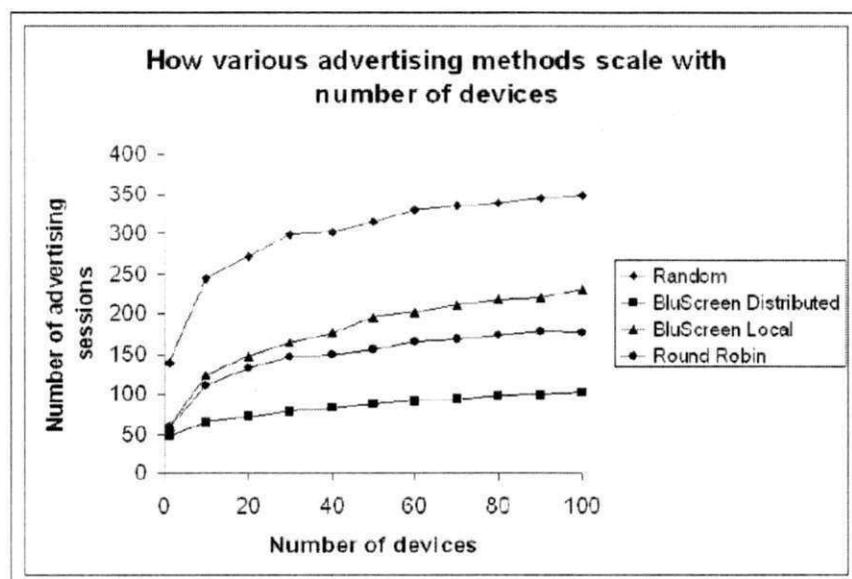


Figura 3.3: Resultados do primeiro experimento. No eixo y está a quantidade de anúncios exibidos e no eixo x a variação na quantidade de dispositivos [39].

creen para pessoas reais. Verificaram-se aspectos relacionados à: relevância e irrelevância dos anúncios exibidos, capacidade da audiência em lembrar o que foi visto e observações gerais sobre o sistema. O experimento foi conduzido por um período de uma semana com oito participantes e utilizou três painéis eletrônicos instalados na Universidade de Southampton.

Um questionário foi aplicado aos participantes com variáveis quantitativas e qualitativas. Os resultados das variáveis quantitativas (Figura 3.4) mostram que em média 76,13% dos anúncios vistos foram lembrados, 53,17% dos anúncios foram considerados relevantes, 32,99% foram considerados irrelevantes e 12,33% dos anúncios lembrados não foram exibidos pelo sistema. Os dados qualitativos referem-se à opinião da audiência em relação ao sistema e relevou importantes fatores externos que podem ter influenciado o experimento. A partir dessas informações é proposta uma metodologia para avaliar a usabilidade de painéis eletrônicos.

Em ambos os experimentos há pontos discutíveis. No primeiro, avaliam-se os algoritmos em relação a sua capacidade de gerar recomendações de anúncios não vistos. Apesar de validar o algoritmo proposto em relação a seu objetivo, essa não é uma informação que comprovadamente aumente a utilidade dos anúncios. Além disso, o uso de uma simulação minimiza a validade do experimento. No segundo experimento esses problemas foram par-

TotalAdverts	% Recall	% Relevant	% Non-relevant	% Extra
13	84.62%	46.15%	30.77%	30.77%
16	62.50%	43.75%	56.25%	0.00%
5	100.00%	60.00%	20.00%	20.00%
7	85.71%	57.14%	28.57%	42.86%
20	80.00%	60.00%	35.00%	0.00%
21	76.19%	33.33%	33.33%	0.00%
20	80.00%	60.00%	30.00%	0.00%
20	40.00%	65.00%	30.00%	5.00%
15.25	76.13%	53.17%	32.99%	12.33%

Figura 3.4: Resultados do segundo experimento relacionados a relevância dos anúncios (*relevant*), capacidade da audiência em lembrar dos anúncios exibidos (*recall*) e anúncios lembrados, mas que não foram exibidos (*extra*) [19].

cialmente eliminados. A avaliação foi realizada com pessoas reais e verificou-se o quanto os algoritmos haviam selecionados anúncios relevantes à audiência. Entretanto o formato como o experimento foi conduzido apresenta alguns problemas. Primeiramente pela pequena quantidade de anúncios que foram usados, apenas 20, e pela pequena quantidade de categorias usadas, apenas três. Esse fato pode induzir um viés, pois pode haver uma alta semelhança entre os anúncios. Outro fato questionável é a pequena amostra de entrevistados que foi utilizada, apenas oito pessoas, todas do sexo masculino. Essa quantidade não pode ser considerada representativa de uma população.

*BluScreen* difere do método proposto nesta dissertação em diferentes aspectos. A principal diferença está no método que os algoritmos usam para selecionar os anúncios considerados relevantes. *BluScreen* considera que, para as pessoas detectadas, quanto menos elas tenham visto o anúncio maiores as chances dela considerar relevante. Outro fator é a necessidade de dispositivos equipados com *Bluetooth* para obter resposta da audiência.

### 3.3 Sistema de recomendação para painéis públicos

O sistema de recomendação para painéis públicos, descrito em [37] tem por objetivo automatizar a seleção automática de conteúdo em sinalização digital. O método proposto utiliza uma abordagem para selecionar o conteúdo balanceando os interesses dos donos dos estabelecimentos onde os painéis estão instalados com os interesses da audiência presente no

local.

O autor argumenta que pelo ambiente onde os painéis estão instalados serem públicos, no sentido de ser uma área de convívio compartilhada, é necessário que haja um equilíbrio entre os interesses da audiência e de quem gerencia o local. O objetivo do método é encontrar conteúdo de interesse a ambos, a partir de diferentes fontes de notícias online. É destacado que o modelo pode ser aplicado a outros cenários, dentre eles o publicitário.

A abordagem é dividida em três subsistemas: perfil do local, selecionador e escalonador. No perfil do local é onde se representam as palavras-chave usadas para representar os interesses da audiência e do estabelecimento, fornecidas por seu gestor. É proposto um modelo de interação onde a audiência indica seus interesses modificando o nome associado ao seu dispositivo *Bluetooth*. O sistema faz uma busca constante por dispositivos *Bluetooth* próximos ao painel e ao identificar algum que tenha seu nome associado a um interesse salva a informação como uma palavra-chave. O gestor do estabelecimento utiliza uma interface própria para inserir suas palavras-chave, gerenciar o painel e contextualizar o ambiente, removendo ambiguidades semânticas que facilitarão a interpretação das palavras-chave. O selecionador é responsável por buscar na Internet conteúdo relevante às palavras-chave definidas no perfil do espaço. Após esse processo é definida uma lista dos assuntos que são adequados à exibição. O escalonador é o responsável por escolher, dentro dessa lista, o conteúdo que será exibido. Essa escolha é baseada em diferentes critérios, onde é feito um balanceamento entre os interesses do dono do estabelecimento e os interesses da audiência. A restrição é que a mensagem seja: *a)* oportunista e sem ter perdido sua importância com o tempo, por exemplo, não exibir o anúncio para um show que já ocorreu; *b)* considerar a popularidade das palavras-chave para determinar se o conteúdo deve ser repetido; *c)* analisar o histórico de recomendação para prevenir que conteúdos similares sejam exibidos e *d)* analisar o conteúdo para identificar se o seu idioma, tamanho e links se adequam ao contexto.

A validação do método foi dividida em três partes. Na primeira verificou-se a capacidade do sistema em escolher conteúdo dentro das restrições temporais, ou seja, selecionar conteúdo que esteja atualizado. Na segunda parte foi analisada a viabilidade da seleção de conteúdo a partir da Internet e por fim foi realizada uma avaliação geral utilizando o sistema por completo. Nessa última etapa foram analisados os pontos de vista da audiência e do gerente do estabelecimento em relação ao sistema.

O experimento foi conduzido por três semanas com um painel eletrônico instalado na entrada principal do Departamento de Sistemas de Informação da Escola Superior de Tecnologia do Instituto Politécnico de Castelo Branco, Portugal. Um questionário foi aplicado a 15 pessoas que utilizaram o sistema e dos resultados obtidos o mais importante a esta dissertação é em relação aos resultados do algoritmo de seleção de conteúdo: para 28,6% dos entrevistados, a relevância do conteúdo foi considerável boa; para 64,3%, foi suficiente; para 7,1%, medíocre; e nenhum dos entrevistados considerou o conteúdo irrelevante.

A abordagem apresentada é diferenciada ao considerar a seleção de conteúdo do ponto de vista de quem é afetado pela mensagem. No entanto, ele requer interações constantes da audiência para seu funcionamento e isso pode não ser viável. Apesar do algoritmo de escalonamento ter tido resultado satisfatório em relação ao percentual de conteúdo avaliado como relevante, o experimento usou uma amostra pequena e isso limita sua capacidade em ser representativo para uma população.

O algoritmo de escalonamento de modo semelhante ao método proposto nessa dissertação utiliza os interesses como fonte para encontrar conteúdo relevante. A diferença entre os dois se dá pelo fato do algoritmo de escalonamento depender de resposta explícita da audiência para seu funcionamento. Além disso, há também a dependência de um dispositivo equipado com *Bluetooth* para que essa interação ocorra.

### 3.4 Targeted advertisement in retail environments

Em [42] é apresentado o desenvolvimento de um *middleware* para fornecer informações de contextos a painéis eletrônicos interconectados em rede. O uso dessas informações possibilita, dentre outras coisas, a seleção de anúncios mais adequados ao contexto da audiência.

O protótipo desenvolvido utiliza uma abordagem distribuída para obter e compartilhar as informações de contexto. Seu desenvolvimento teve como base um framework de gerenciamento de contexto denominado *Context Management Framework* (CMF). O CMF usa uma abordagem multiagentes para obter, processar e armazenar informações de contextos obtidas por diferentes sensores. Com isso facilita-se sua distribuição aos diversos painéis, possibilitando que em cada painel seja definido seu próprio contexto.

Para demonstrar a viabilidade do projeto foram desenvolvidos dois estudos de caso. Ape-

nas o estudo de caso relacionado à seleção de anúncios será discutido. Inicialmente esse mesmo estudo de caso foi proposto em [41], porém apresenta-se uma versão aprimorada. Sua proposta é selecionar anúncios mais relevantes aos interesses da audiência próxima ao painel.

O sistema desenvolvido utiliza sensores *Bluetooth* para se comunicar com a audiência próxima ao painel eletrônico. Para minimizar problemas relacionados à privacidade, o rastreamento apenas identifica dispositivos que foram previamente cadastrados no sistema. Nesse processo de cadastro os usuários registram o *MAC* de seu dispositivo juntamente com seus interesses. Ao identificar dispositivos cadastrados o sistema usa esses interesses para exibir anúncios relacionados a eles. Junto ao anúncio há um código QR que fornece mais informações sobre o produto anunciado, como a localização da loja que o comercializa. Esse código também é usado como forma de obter resposta da audiência, sendo considerado um indício de que o usuário teve interesse pelo que foi anunciado. Uma ilustração do cenário e design da arquitetura é apresentada na Figura 3.5.

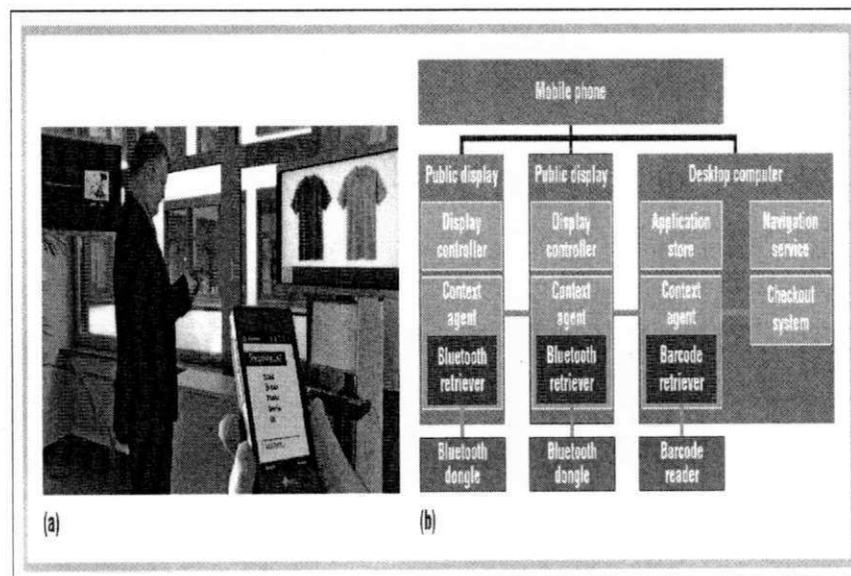


Figura 3.5: a) exemplo de aplicação do sistema; b) arquitetura do sistema [39].

Não são fornecidos detalhes técnicos sobre como o algoritmo seleciona os anúncios, apenas é informado o dado que é utilizado nessa seleção.

O ponto negativo do estudo é que não é realizado um experimento para avaliar a viabilidade do *middleware* em selecionar conteúdo relevante aos interesses do usuário. Além disso,

a necessidade em cadastrar os dispositivos *Bluetooth* e requisitar explicitamente as preferências dos usuários é outro fator considerado negativo. Em comparação ao algoritmo proposto nesta dissertação, essas são as principais diferenças entre os métodos.

### 3.5 Advertisement Recommendation Algorithm (ARA)

Em [46] é apresentada uma infraestrutura para seleção de publicidade em painéis eletrônicos de sinalização digital móveis. O fato de considerar painéis eletrônicos móveis, por exemplo, instalados em carros ou ônibus, é um diferencial dessa abordagem. Os anúncios são selecionados com base na localização geográfica atual do painel e os horários mais adequados a sua exibição para uma determinada localidade.

A infraestrutura é dividida em módulos que são responsáveis pelo gerenciamento dos painéis, anúncios e por sua recomendação. É utilizada uma arquitetura centralizada conectada à Internet, onde os painéis enviam a ela dados sobre sua localização geográfica e requisitam os anúncios para serem exibidos. Os critérios para seleção dos anúncios são definidos pelos anunciantes, que o fazem através de uma interface web. Os módulos da infraestrutura são listados na Figura 3.6.

O mecanismo de recomendação dos anúncios, denominado *Advertising Recommendation Algorithm (ARA)*, baseia suas recomendações na localização geográfica dos painéis, obtidas por meio de um GPS, e na informação sobre qual o melhor horário para exibir os anúncios. Para esse cálculo utiliza-se a lógica nebulosa que determinará a relevância com base no nível de pertinência do horário atual em relação à função que define os melhores horários para exibição do anúncio. O cálculo retorna um dos cinco valores de relevância definidos, como: relevância absoluta, alta relevância, relevância média, pouca relevância e irrelevante.

Tendo como objetivo validar o método um experimento foi simulado durante um período de 30 dias entre as 6 da manhã a 10 da noite. Seu objetivo foi verificar quais anúncios foram recomendados enquanto os painéis estavam em determinada localização geográfica. O algoritmo ARA foi comparado a outros três métodos de seleção de anúncios: seleção por localização, sequencial e aleatória. Na seleção por localização o anúncio é exibido quando o painel alcança determinada região. Na sequencial os anúncios são recomendados seguindo uma ordem fixa e na aleatória usa-se a aleatoriedade para selecionar os anúncios. Para si-

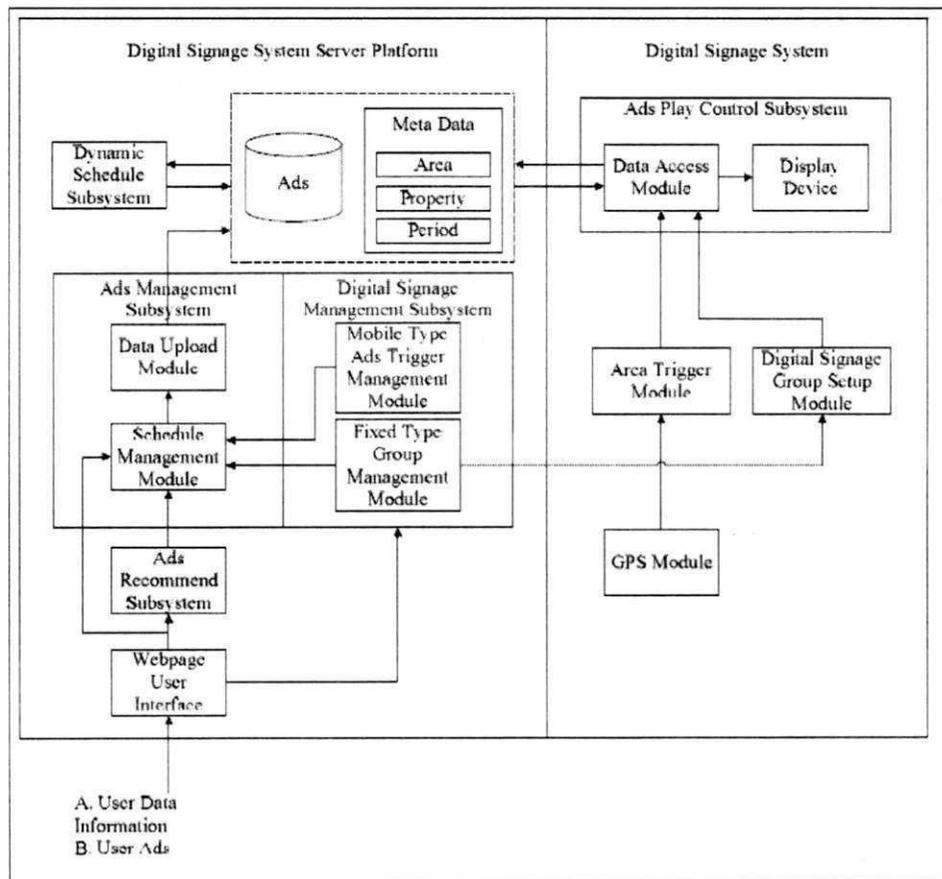


Figura 3.6: Arquitetura do sistema de sinalização digital [46].

mular o experimento foram usados dados geográficos da cidade de Hsinchu, localizada em Taiwan, obtidos pelo *Google Maps*. A área da cidade foi dividida em regiões classificadas como: educacionais, comércio, médica, residencial e turística. Os anúncios a serem exibidos foram classificados em 7 categorias e, para cada categoria, definiram-se os melhores horários para exibí-los. Com base no mapa de Hsinchu foi simulado o tráfego de um veículo portando um painel eletrônico e trafegando a uma velocidade de 10km/h. O painel exibia anúncios com 3 segundos de duração.

O objetivo do experimento foi verificar quais anúncios eram selecionados quando o veículo estava em uma determinada região. Pelos resultados os autores concluíram que o algoritmo ARA selecionou os anúncios mais adequados para as regiões por onde o veículo transitou.

A execução do experimento é passível de contestação em alguns pontos. Verificar se o

algoritmo seleciona determinado tipo de anúncio para uma região não significa que ele será de interesse à audiência, caso não haja uma comprovação empírica de que esse método se aplica, por exemplo, entrevistando a audiência desses locais, fato que o trabalho não o faz por ser simulado.

A similaridade entre o algoritmo ARA e o método proposto nesta dissertação é que ambos não necessitam de dados explícitos da audiência. No entanto, eles diferem em relação aos dados necessários ao funcionamento dos algoritmos. Outro diferencial é a necessidade de um GPS para o funcionamento do algoritmo ARA.

### 3.6 MyAds

O sistema *MyAds* proposto em [9] tem por objetivo melhorar a seleção de anúncios em sinalização digital, fazendo a seleção dos anúncios mais adequados aos interesses da audiência. Para esse processo de seleção o sistema depende de dados da audiência que são coletados por meio de sensores, como microfones, câmeras, dispositivos móveis, entre outros.

*MyAds* é composto por três módulos: *Content display*, *Information processing* e *Content allocation*. Esses módulos são responsáveis pelo gerenciamento dos painéis, coleta dos dados da audiência, inferências do seu perfil e adaptação do conteúdo do painel ao que foi inferido. O fluxo dessas etapas está representado na Figura 3.7.

Como pode ser visualizado no gráfico, o fluxo inicia com a obtenção do perfil da audiência (1). Esse perfil é formado por dados que os usuários informam explicitamente, mas também pode ser agregada informação de contexto obtida por meio de diversos sensores como câmeras, microfones, dentre outros. Em seguida o perfil é transmitido para que sejam inferidos conhecimentos usáveis pelo sistema (2a). De posse desse conhecimento (3) o sistema oferece espaço publicitário por meio de um mecanismo de lances. Os lances são feitos pelos anunciantes com base na importância do conhecimento inferido pelo perfil da audiência aos anúncios.

O perfil da audiência é parte fundamental desse processo, ele é formado por dados demográficos como idade e gênero, e pelo interesse por algumas categorias. Para determinar o melhor anúncio para ser exibido é feita uma associação entre esses interesses aos interesses relacionados ao anúncio em um processo de *matching*. O algoritmo verifica apenas se há

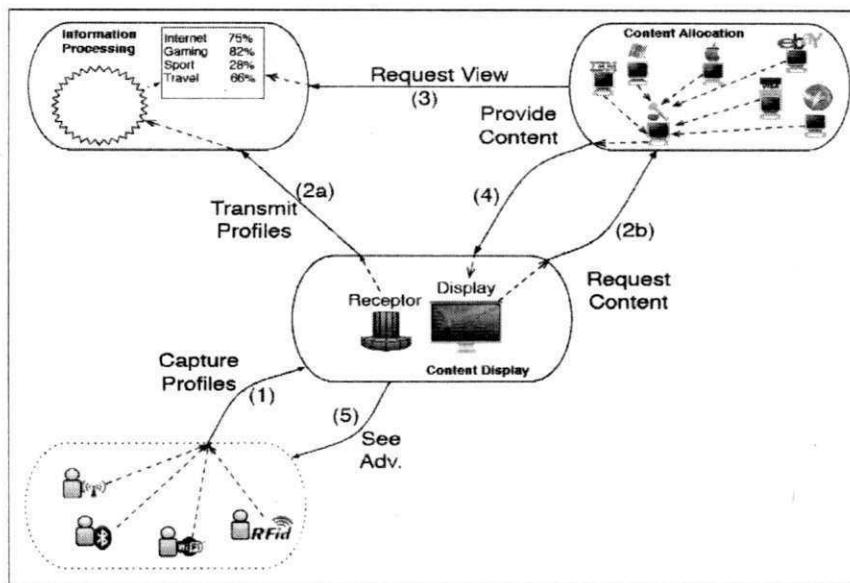


Figura 3.7: Arquitetura MyAds [9].

ou não determinado interesse no perfil do usuário, quanto maior essa quantidade, maior a chance do anúncio ser relevante.

Para validar o trabalho três abordagens para seleção de anúncios foram comparadas. Duas são propostas pelo autor, nomeadas como *Best-fit* e *Auction*. No primeiro modelo o sistema seleciona o anúncio com maior probabilidade de ser relevante ao perfil da audiência. Nesse caso os anunciantes definem o perfil das possíveis pessoas interessadas pelo seu anúncio e o algoritmo seleciona para exibição o que considerar mais relevante. No modelo *Auction* é utilizada uma abordagem parecida, no entanto é usado um mecanismo de leilão onde os anunciantes fazem lances com base em quanto eles consideram o perfil da audiência relacionado aos seus anúncios. O método comparado para validar o trabalho, denominado *First Come First Served* (FCFS), é usado por sistemas tradicionais de sinalização digital e consiste na seleção de anúncios usando um modelo de reserva prévia de espaço publicitário, semelhante aos moldes dos anúncios para televisão, onde os anunciantes pagam por um determinado espaço para exibir seu anúncio por um período pré-definido. Nesse caso o perfil da audiência não exerce qualquer influência.

Um experimento simulado foi conduzido para validar as propostas. Seus objetivos foram: *i)* avaliar a relevância dos anúncios selecionados e *ii)* qual das propostas é mais eficiente em relação a economizar os recursos financeiros dos anunciantes, visto que há o pagamento

pelo espaço para anunciar. Pelo contexto desta dissertação apenas o primeiro objetivo foi analisado.

Dois cenários foram usados para validar o estudo. No primeiro foram simulados os perfis de usuários, onde os seus interesses foram selecionados aleatoriamente com a mesma probabilidade de serem escolhidos, havendo com isso uma heterogeneidade nos interesses da amostra. No segundo cenário houve a priorização de determinado interesse, para que houvesse um interesse dominante na amostra. Em ambos os cenários foi verificada a quantidade de anúncios selecionados que tinham relação com os interesses dos usuários, em um processo chamado pelo autor de *matching*. Os resultados para o primeiro e segundo cenário estão listados nos gráficos das Figuras 3.8 e 3.9, respectivamente.

A parte azul das barras indica o percentual de anúncios selecionados e que tinham relação com o perfil da audiência. Na parte vermelha é indicado o percentual de anúncios selecionados, mas sem relação. Os resultados mostram que em ambos os cenários os algoritmos propostos são melhores que a abordagem FCFS. No entanto, no segundo cenário onde os interesses são menos individualizados o resultado foi mais evidente.

Há dois pontos negativos no experimento. O primeiro é o fato de ele ser simulado e analisar o desempenho dos algoritmos apenas em relação a sua capacidade em encontrar anúncios relacionados ao perfil da audiência. Isso não evidencia a capacidade do algoritmo em selecionar os melhores anúncios para a audiência. Como consequência, pela ausência de pessoas participando dos testes não foi avaliada empiricamente a capacidade dos algoritmos em recomendar anúncios relevantes.

Apesar da descrição do perfil da audiência usar dados parecidos aos utilizados pelo algoritmo proposto nesta dissertação, esse trabalho se diferencia principalmente pela forma como o problema é solucionado e pelo fato de não usar dados explícitos da audiência nem informações de contexto.

### 3.7 Comparação entre os métodos

Cada um dos métodos apresentados possuem suas vantagens e desvantagens dependendo do cenário onde sejam aplicados. Pela casualidade do contato entre os consumidores e a sinalização digital, obter informações dessas pessoas explicitamente pode não ocorrer. Além

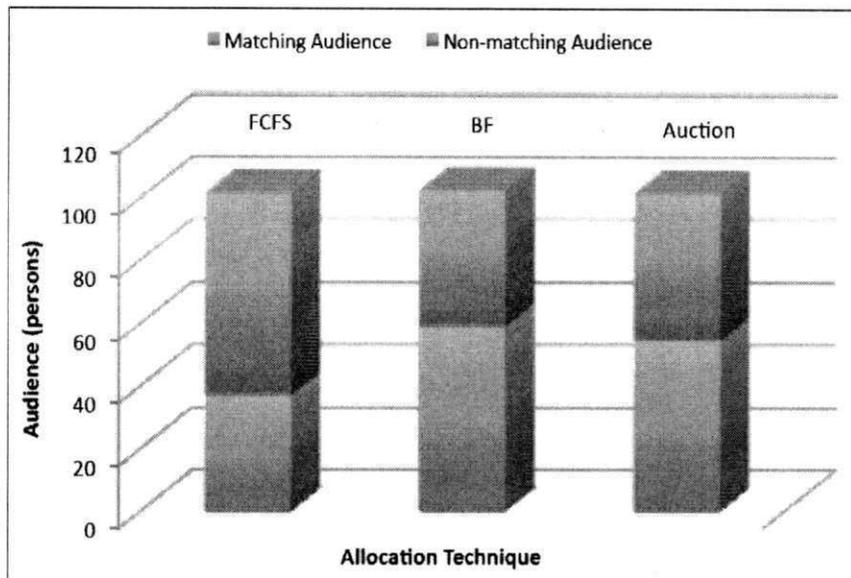


Figura 3.8: Resultados do experimento realizado para o primeiro cenário [9].

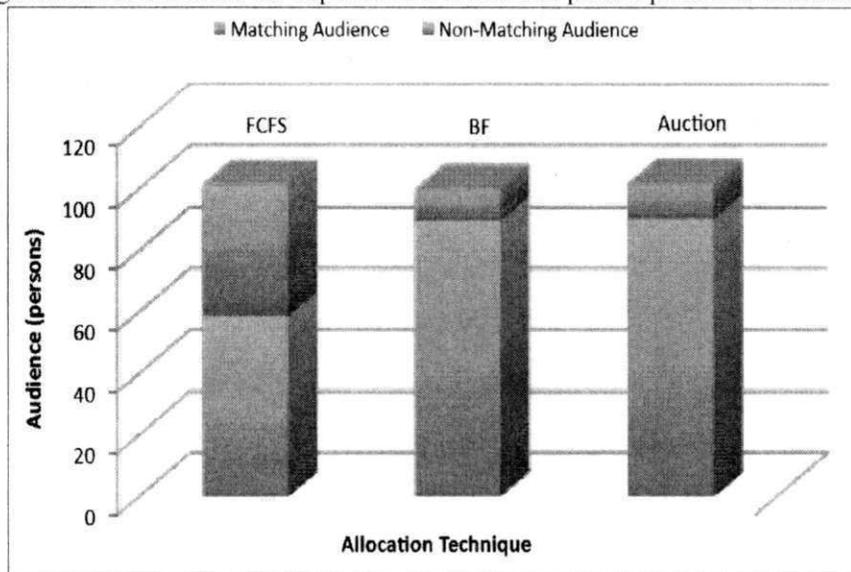


Figura 3.9: Resultados do experimento realizado para o segundo cenário [9].

disso, é um processo de difícil escalabilidade, diante da grande variedade de sistemas existente na sinalização digital [3]. Por outro lado, o uso de dispositivos tecnológicos pode impor uma barreira ao uso do método e possivelmente aumentar os custos do projeto.

Diferente dos métodos apresentados, o método proposto nesta dissertação não necessita de interação com a audiência ou de dispositivos tecnológicos. Sua limitação está na necessidade de conhecimento especialista para definir os anúncios e estabelecimentos, porém essa é uma informação que pode ser obtida de diversas fontes, como publicitários, proprietário dos

estabelecimentos, agências de publicidade, empresas que comercializam os produtos anunciados, dentre outros. Além disso, a quantidade de informações que precisam ser definidas é relativamente pequena, o que além de tornar a solução escalável não impossibilita sua definição.

Na Tabela 3.1 é apresentado um comparativo entre o método proposto e os demais apresentados nessa seção. É relacionado o tipo de entrada que essas abordagens utilizam, necessidade de interação com a audiência, dispositivos tecnológicos ou de conhecimento especialista.

Método	Solução	Forma de entrada	Necessita de interação com a audiência?	Depende de dispositivos tecnológicos (quais)?	Necessita de conhecimento especialista?
Método proposto	Seleção baseada no uso combinado dos interesses e da informação do público alvo dos produtos	Implícito	Não	Não	Sim
<i>Mobidic</i>	Seleção baseada em dados de contexto	Explícito	Sim	Sim (celular)	Não
Sistema de recomendação para painéis públicos	Seleção baseada no uso de palavras-chave	Explícito	Sim	Sim (celular)	Não
<i>Bluscreen</i>	Seleção baseada em probabilidades	Implícito	Não	Sim (aparelhos com dispositivo Bluetooth)	Não
<i>MyAds</i>	Seleção baseada na associação dos interesses dos consumidores e dos anúncios	Explícito	Sim	Sim (variados sensores, dentre eles RFID)	Não
<i>Targeted advertisement in retail environments</i>	Não informado	Explícito	Sim	Sim (necessita de dispositivos com tecnologia Bluetooth)	Não
<i>Advertisement Recommendation Algorithm (ARA)</i>	Lógica nebulosa para encontrar o melhor horário para exibir o anúncio	Implícito	Não	Sim (necessita de GPS)	Sim

Tabela 3.1: Tabela comparativa dos métodos.

## Capítulo 4

# Método para seleção de anúncios em sinalização digital

### 4.1 Introdução

O cenário atual de publicidade requer a exibição de anúncios personalizados para captar mais efetivamente a atenção da audiência, diante da exposição frequente que o consumidor está a inúmeros anúncios em diferentes meios de comunicação. Porém, um dos grandes desafios na busca por essa adequação consiste no desenvolvimento de métodos que melhorem essa seleção sem a necessidade de interação explícita com os consumidores para obter suas preferências ou dependência de dispositivos tecnológicos adaptados ou desenvolvidos para essa finalidade.

Para tornar a seleção de anúncios em sinalização digital mais efetiva, satisfazendo as restrições de não depender de interação com a audiência ou dispositivos tecnológicos, propõe-se um método para seleção de anúncios baseado em conceitos e elementos da teoria do marketing. Para minimizar os problemas citados, as entradas do método serão obtidas implicitamente pela descrição dos interesses e do público-alvo relacionado ao estabelecimento onde os consumidores se encontram. A presença dos consumidores nesse estabelecimento é um indício de suas preferências e essa informação pode ser usada para tornar a seleção de anúncios mais efetiva. No cenário de sinalização digital essa é uma informação conhecida, pois é a mesma dos painéis eletrônicos que possuem uma localização fixa e determinada a priori.

Como descrito em [30], sabe-se que as aquisições do consumidor refletem seus interes-

ses e necessidades. Essas aquisições podem trazer outros indícios, como as características do consumidor. Sabe-se que os consumidores de um determinado produto são, em geral, homogêneos em relação as suas preferências, classe social, faixa etária, gênero ou até mesmo estado civil. Por exemplo, a maior parte das pessoas que visitam lojas da *Carmem Steffens* são mulheres, entre 15 a 45 anos e das classes sociais A e B<sup>1</sup>. Essa informação pode ser usada para alcançar o nível de personalização desejado para os anúncios, recomendando aqueles que tenham relação com os interesses e ao público-alvo desses estabelecimentos [11].

O público-alvo de um anúncio são as pessoas com potencial interesse pelo produto anunciado. Essas pessoas, em geral, compartilham algumas características e prezam por valores similares. Esse formato de comercialização se torna muito efetivo, quando comparado à forma mais tradicional de mercado de massa. Em uma sociedade mais individualizada e com necessidades diferentes, fabricar um mesmo produto a todos não é mais uma alternativa viável para a maior parte das organizações [30]. Através do público-alvo a organização consegue limitar seu escopo de ação, possibilitando identificar com maior precisão as necessidades dos consumidores. Com isso é possível alcançar mais efetivamente as pessoas com maior chance de se interessarem pelo seu produto.

Para complementar o uso do público-alvo na recomendação de anúncios, propõe-se utilizar também os interesses dos consumidores por ser uma informação que fornece mais detalhes sobre suas preferências de consumo [31]. Assim como o público-alvo, os interesses podem ser inferidos pelo local frequentado pelos consumidores. Por exemplo, a presença de um grupo de pessoas em um restaurante de culinária vegetariana é um forte indício do interesse dessas pessoas por alimentos vegetarianos e pelos valores praticados pelos adeptos desse tipo de culinária. Consequentemente, um anúncio relacionado a esse interesse, como o de um congresso sobre vegetarianismo, teria uma maior chance de ser considerado relevante.

Pelo fato das informações que descrevem os estabelecimentos e seu público-alvo fazerem parte de conceitos de marketing bem definidos e ser uma informação de fácil acesso de seus proprietários, isso dá credibilidade ao seu uso no processo de recomendação de anúncios.

Com base nas informações descritas anteriormente são propostos três métodos para recomendação dos anúncios:

- *Método de seleção baseado na similaridade de interesses*: seu objetivo é selecionar

---

<sup>1</sup><http://www.pratike.com.br/Noticias.asp?autonumero=1>

anúncios com interesses mais similares aos interesses do estabelecimento onde se encontra o grupo de consumidores.

- *Método de seleção baseado na proximidade entre o público-alvo dos anúncios e estabelecimento*: o anúncio com público-alvo classificado como mais próximo ao público-alvo do estabelecimento será selecionado.
- *Hibridização dos métodos*: as abordagens anteriores serão agregadas, de modo a minimizar as limitações que esses métodos teriam se usados em separado e aproveitar suas vantagens individuais.

## 4.2 Modelo para representação dos anúncios e estabelecimentos

Para garantir que os conceitos propostos sejam bem compreendidos, nessa seção será apresentada a formalização matemática dos elementos relacionados ao modelo de recomendação de anúncios. O formalismo adotado é baseado na teoria dos conjuntos, pois seus fundamentos são adequados à expressão formal desse problema.

**DEFINIÇÃO 4.2.1:** Anúncio publicitário -  $A$ . Forma de comunicação utilizada pelas empresas para a divulgação de produtos aos seus potenciais consumidores. A essas pessoas é dado o nome de público-alvo. Define-se matematicamente um anúncio como:

$$A = \langle P, PA \rangle$$

Onde:

- $P$  é o produto anunciado.
- $PA$  é o público-alvo ao qual o anúncio é direcionado.

**DEFINIÇÃO 4.2.2:** Produto -  $P$ . Para o marketing um produto pode ser um bem tangível, serviço ou uma marca. Ele é oferecido ao mercado para consumo e direcionado a satisfazer necessidades ou desejos [31].

As necessidades e desejos aos quais o produto satisfaz estão relacionados a interesses. Esses interesses refletem assuntos com alguma ligação com o produto. Como exemplo de relação, pode-se definir que um vídeo game remete à entretenimento e jogos.

Um produto pode estar relacionado a uma ou mais categorias de interesses. Por exemplo, uma academia pode ter associada a ela vários interesses distintos, tais como alimentos saudáveis, roupas para academia, suplementos alimentares e eventos esportivos. Define-se um produto como:

$$P = \langle Desc, I \rangle$$

Onde:

- Desc é uma descrição textual sobre o produto anunciado;
- I é o conjunto de interesses relacionados ao produto.

**DEFINIÇÃO 4.2.3:** Público-alvo - *PA*. O público-alvo de um produto é representado pelas características presentes na maioria dos consumidores de determinado produto. Essas características variam em sua forma, podendo ser demográfica, psicográfica, comportamental ou geográfica. Nesta dissertação a representação do público-alvo utiliza apenas os dados demográficos dos consumidores por ser a informação mais utilizada nesse processo [30], além de ser um dado viável de obter.

A definição do público-alvo baseia-se no modelo proposto em [30]. São usados o gênero, faixa etária e classe social. O gênero indica para qual sexo o produto é direcionado, havendo três possibilidades: masculino, feminino ou unissex. A faixa etária é uma faixa de valores que representa a idade majoritária dos consumidores do produto, enquanto a classe social indica quais as classes sociais com maior possibilidade de adquirir o produto. Os valores aceitos para essa variável são definidos pelo critério de classificação econômica Brasil [1].

Uma pessoa pertence ao público-alvo de um produto, caso seus dados demográficos estejam presentes na faixa de valores aceitáveis para esse público-alvo. Essa faixa de valores indica os valores aceitos para cada uma das variáveis: gênero, faixa etária e classe social. Para cada valor aceito é associado um nível de certeza que representa o quanto ele pertence ao público-alvo. Isso possibilita a representação de valores com uma menor certeza em per-

tencer ao público-alvo, mas que são relevantes. Com base nessas informações propõe-se a seguinte formalização:

$$PA = \langle G, CS, FE \rangle$$

Onde:

$$G = \{(S, \phi) | S \in \{M, F\} \wedge \phi \in \mathbb{Q} | 0 \leq \phi \leq 1\}$$

$$CS = \{(Cs, \phi) | Cs \subset \{A1, A2, B1, B2, C1\} \wedge \phi \in \mathbb{Q} | 0 \leq \phi \leq 1\}$$

$$FE = \{(intervalo, \phi) | intervalo \subset \mathbb{N}^* \wedge \phi \in \mathbb{Q} | 0 \leq \phi \leq 1\}$$

Onde:

- G é o conjunto de gêneros pertencentes ao público-alvo do produto;
- $\phi$  é um valor entre 0 e 1 que indica o nível de certeza de quanto determinado valor pertence ao público-alvo. O valor 0 indica que não há relação, enquanto o valor 1 indica uma forte relação.
- CS é o conjunto das classes sociais aceitas para o público-alvo do produto;
- FE é o conjunto de faixa etárias ao qual o produto é direcionado.
- Intervalo representa uma faixa etária formada por duas idades que limitam sua ação. Onde x representa um valor possível para a faixa etária. Sua representação seguirá o padrão ISO 31-11:

$$]a, b[ = \{x \wedge a < x < b\}$$

$$[a, b[ = \{x \wedge a \leq x < b\}$$

$$]a, b] = \{x \wedge a < x \leq b\}$$

$$[a, b] = \{x \wedge a \leq x \leq b\}$$

**DEFINIÇÃO 4.2.4:** Interesse -  $I$ . Representa uma categoria de assunto ao qual o produto tem ligação. Essa ligação pode ser direta, como exemplo, a relação entre vídeo games e entretenimento, ou indireta, como um calçado a festas. Interesses são representados pelo nome da categoria de assunto que ele representa:

$$I = \langle \textit{nome} \rangle$$

Onde:

- nome: representa o nome da categoria de interesse, por exemplo: entretenimento.

**DEFINIÇÃO 4.2.5:** Estabelecimento -  $E$ . Local que oferta algum serviço a um determinado público-alvo. Assim como o produto, ele está relacionado a uma ou mais categorias de interesses. Sua definição é a seguinte:

$$E = \langle \textit{Nome}, I, PA \rangle$$

Onde:

- Nome é uma descrição textual sobre o estabelecimento;
- $I$  é o conjunto de interesses relacionados ao estabelecimento;
- PA é o público-alvo ao qual o estabelecimento é direcionado.

### 4.3 Método de recomendação baseado na similaridade de interesses

Os interesses do consumidor possuem uma forte influência em suas decisões. Os produtos adquiridos, locais visitados ou programas de televisão assistidos, são exemplos de decisões motivadas pelos interesses do consumidor. Exibir anúncios de produtos que tenham relação com esses interesses é uma alternativa para aumentar a efetividade do anúncio.

O método proposto baseia as recomendações de anúncios na similaridade entre as categorias de interesses associadas ao estabelecimento e aos produtos anunciados. Pode haver a

relação com uma ou mais categorias de interesses. Como exemplo, a marca de óculos Chilli Beans tem seus produtos associados a esportes, música e moda <sup>2</sup>, enquanto um hotel pode estar relacionado a turismo e entretenimento. Essa informação pode indicar uma variedade de características dos consumidores, como seus hábitos alimentares, esportes praticados, estilo musical, preferência por atividades de diversão, dentre outras informações. A definição das categorias de interesses e sua relação com os produtos e estabelecimentos podem ser realizadas por um especialista, como um publicitário.

Para representar esses interesses serão utilizados conjuntos matemáticos, como definido na seção 4.2. A definição das categorias de interesses relacionadas aos produtos e estabelecimentos pode ser realizada por um especialista, como exemplo um publicitário. Por exemplo, os interesses relacionados à loja de departamento C&A e a um anúncio de máquinas fotográficas, podem ser definidos da seguinte forma:

$$I_{cea} = \{ \text{entretenimento, moda, beleza, sapatos, vestuário} \}$$

$$I_{\text{máquina fotográfica}} = \{ \text{fotografia, entretenimento, turismo, arte, filmagem} \}$$

Para determinar a similaridade entre os interesses será utilizado o índice de *Jaccard*. Seu cálculo consiste em dividir o número de elementos comuns aos conjuntos pela união dos seus elementos. O cálculo é determinado pela seguinte fórmula:

$$\text{sim}(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

A função  $\text{sim}(A, B)$  calcula a similaridade que há entre os conjuntos A e B e retorna um valor dentro de uma faixa de 0 a 1. O valor 0 indica a completa ausência de elementos em comum aos conjuntos e 1 representa a semelhança completa.

Para exemplificar o cálculo, considere as representações feitas acima para os interesses relacionados ao estabelecimento C&A e ao anúncio de uma máquina fotográfica. Havendo mais um anúncio relacionado a grife americana Victoria's Secret, qual anúncio seria mais adequado a exibição nesse estabelecimento? A representação dos interesses relacionados a Victoria's Secret é:

<sup>2</sup><http://www.doisnovemeia.com.br/almanaque/documentos/briefingConcurso.pdf>

$$I_{victoriasecret} = \{moda, grife, beleza, vestuário\}$$

O cálculo do índice de *Jaccard* entre os interesses da C&A e dos anúncios é representado abaixo:

$$sim(I_{cea}, I_{máquina\ fotográfica}) = \frac{|A \cap B|}{|A \cup B|} = \frac{1}{9} = 0.11$$

$$sim(I_{cea}, I_{victoriasecret}) = \frac{|A \cap B|}{|A \cup B|} = \frac{3}{5} = 0.60$$

Como previsto, por apresentar mais categorias de interesses similares ao estabelecimento, o anúncio da Victoria's Secret obteve um índice de *Jaccard* maior que o anúncio da máquina fotográfica, 0.60 contra 0.11.

O seguinte algoritmo é proposto para calcular a similaridade entre os conjuntos e identificar o melhor anúncio:

**ALGORITMO PARA O MÉTODO DE SELEÇÃO DE ANÚNCIOS PELA SIMILARIDADE DOS INTERESSES:**

1. Dado um conjunto de interesses associados ao estabelecimento
2. Para cada anuncio:
  - Calcular a similaridade entre os interesses.
3. Ordenar os anúncios de forma crescente em relação ao valor de similaridade obtido.
4. Verifica se o anúncio com maior similaridade já foi exibido.
  - Se a condição for verdadeira, repete o procedimento para o anúncio seguinte.
  - Se a condição não for verdadeira, seleciona o anúncio para exibição.

## 4.4 Método de seleção baseado na proximidade entre o público-alvo dos anúncios e do estabelecimento

Em um mercado cada vez mais individualizado, o número de organizações que produzem produtos sem personalização para diferentes mercados é reduzido. A tática de mercado de

massa, anteriormente efetiva, dá lugar à segmentação de mercado. O princípio dessa prática consiste na divisão do mercado para que as ações de marketing foquem em grupos menores de consumidores, com necessidades mais definidas.

As pessoas que fazem parte desses grupos e que são de interesse de uma organização recebem o nome de público-alvo. O público-alvo é definido a partir de análises de mercado que identificam as características comuns à maioria dos consumidores de determinado produto. As pessoas que possuem essas características têm maiores chances de se interessar pelo produto, pois compartilham princípios e valores semelhantes. Com essa informação a organização pode identificar mais efetivamente as necessidades, hábitos e canais de comunicação que esses consumidores utilizam, para que seus esforços de marketing, por exemplo, a divulgação de um anúncio, seja direcionada aos meios de comunicação usados por essas pessoas. Como exemplo, os gestores de marketing da Lindt, fabricante de chocolates suíços, têm conhecimento de que 80% do consumo de seu chocolate é feito por mulheres de 20 aos 55 anos. Essa informação é a base para a promoção da empresa que direciona seus esforços aos canais de comunicação a que essas mulheres têm contato<sup>3</sup>. Da mesma forma os estabelecimentos são frequentados por pessoas que possuem certo grau de semelhança em suas características. Por exemplo, a franquia de lanches saudáveis Subway sabe que a maioria dos seus consumidores é formada por pessoas de 16 a 39 anos<sup>4</sup>.

Propõe-se usar as características que descrevem o público-alvo de um estabelecimento para selecionar os anúncios direcionados a essas características. A presença das pessoas nesse local é um indício de que elas possivelmente possuem alguma característica de seu público-alvo. Conhecendo o público-alvo dos anúncios é possível identificar qual anúncio é direcionado ao público-alvo do estabelecimento. Por exemplo, há dois anúncios de viagens para o carnaval, o primeiro é para um cruzeiro marítimo de 8 noites e o segundo é para passar o carnaval em Recife em um hotel econômico (Figura 4.1). Em um restaurante com público-alvo direcionado a pessoas de classe social A1 e A2, possivelmente o anúncio do cruzeiro seria mais relevante. Já em um restaurante com público-alvo formado por pessoas da classe econômica B2 e C, o anúncio do carnaval em Olinda parece ser mais adequado. Nesse exemplo, o fator financeiro limita ou possibilita o acesso ao produto anunciado.

<sup>3</sup><http://kanada.ahk.de/index.php?id=9507>

<sup>4</sup><http://www.subway.co.uk/business/franchise/faqs.aspx>



DEMOGRÁFICOS	
Idade	0-6, 6-11, 12-19, 20-34, 35-49, 50-64, 65+
Sexo	Masculino, Feminino
Família	1 ou 2 pessoas, 3 ou 4 pessoas, 5 pessoas ou mais
Renda	0-\$10,000; \$10,000-\$20,000; \$20,000-\$30,000; \$30,000-\$50,000; \$50,000+
Ocupação	Profissional, técnico, gerente, oficial, vendedor, desempregado, etc.
Educação	Analfabeto, Ensino Fundamental, Ensino Médio, Ensino Superior, Pós-Graduado
Religião	Católico, Protestante, Judeu, Muçulmano, Hindu, entre outras
Raça	Asiáticos, Hispânicos, Negros, Brancos
Nacionalidade	Norte-Americanos, Sul-Americanos, Britânicos, Franceses, Italiano, Japoneses
PSICOGRÁFICOS	
Classe Social	A1, A2, B1, B2, C1, C2, D, E
Estilo de vida	Realizadores, Esforçados, Sobreviventes
Personalidade	Compulsiva, Autoritária, Ambiciosa
GEOGRÁFICOS	
Região Mundial	América do Norte, Oeste Europeu, Ásia
País	Canadá, Alemanha, China, México, Brasil
Densidade	Urbano, Suburbano, Rural
Clima	Tropical, Subtropical, Equatorial, Mediterrâneo, etc.
COMPORTAMENTAL	
Ocasões	Ocasão Regular; Ocasão Especial
Benefícios	Qualidade, Serviço, Economia, Conveniência, Velocidade
Status do usuário	não-usuário, ex-usuário, usuário em potencial, primeira vez, usuário regular
Lealdade	Nenhuma, Média, Forte, Absoluta
Conhecimento	Não Conhece o Produto, Consciente, Informado, Interessado, Desejoso
Atitude em relação ao produto	Entusiasta, Positiva, Indiferente, Negativo, Hostil

Tabela 4.1: Variáveis usadas no processo de segmentação de mercado.

sentado com base no critério de classificação econômica Brasil [1], reduzindo a necessidade de especificar valores detalhados. Para a idade serão usados intervalos de valores para indicar a faixa etária que descreve os consumidores de um produto.

Para exemplificar sua aplicação, considerando a formalização do público-alvo descrita na seção 4.2, o público-alvo para um anúncio da marca de roupa íntima Victoria's Secret pode ser definido como [33]:

$$PA_{victoriasscret} = \langle G_{vs}, FE_{vs}, CS_{vs} \rangle$$

$$G_{vs} = \{(F, 1)\} \text{ Direcionado a pessoas do sexo feminino.}$$

$$FE_{vs} = \{([15, 35], 1)\} \text{ Mais utilizado por mulheres entre 15 e 35 anos.}$$

$$CS_{vs} = \{(A1, A2, B1, 1)\} \text{ Para mulheres da classe social A e B.}$$

No entanto, essa definição apresenta um inconveniente. Como tratar o público-alvo de um estabelecimento  $PA_{estabelecimento} = \langle \{(F, 1)\}, \{([40, 50], 1)\}, \{(A1, 1)\} \rangle$ , formado por mulheres de 40 a 50 anos, faixa etária ausente do público-alvo, mas de classe social presente nele. Considerando o exemplo, percebe-se que o público do estabelecimento é bastante similar ao público-alvo do anúncio, exceto pela idade. Apesar dessa diferença, faz sentido ponderar que quanto mais próximas as características do público-alvo do estabelecimento forem do público-alvo de um produto, maior é a chance do produto ser relevante.

O exemplo acima é uma demonstração de que o processo de classificar um conjunto de entradas como pertencentes ou não ao público-alvo de um anúncio não é um caso de lógica clássica, pois deve possibilitar margens para aceitar valores próximos [30]. Sua solução requer que além da definição do público-alvo dos anúncios e estabelecimentos, também sejam definidos níveis para aceitar valores próximos. Quanto maior a proximidade entre esses valores, maior são as chances das pessoas com essas características fazerem parte do público-alvo e a mensagem do anúncio ser considerada relevante.

A lógica nebulosa apresenta condições adequadas para resolver problemas que envolvem diferentes níveis de aceitação para uma variável, como é a situação. Esses níveis de aceitação podem ser considerados análogos ao uso das funções de pertinência. O uso dessas funções associado ao raciocínio nebuloso possibilitará inferir a relevância do anúncio com base nas proximidades das variáveis de entradas em relação aos valores do público-alvo do anúncio.

Especialistas irão definir as funções de pertinência para mapear as entradas em relação às variáveis idade e rendimento. O gênero não será considerado no raciocínio nebuloso por não haverem níveis de pertinência relacionados aos seus valores. Essa informação será usada para filtrar os anúncios que sejam mais adequados a determinado gênero sexual.

Para indicar a relação entre os valores de entrada e sua relação com as variáveis foram definidos três termos linguísticos: *pertence*, *próximo* ou *ausente*. O primeiro termo indica que o valor *pertence* à variável do público-alvo, ou seja, está dentro da faixa de valores aceitáveis, enquanto o segundo indica que há uma proximidade entre eles e o terceiro termo indica a inexistência de relação. O uso do termo *ausente* possibilita limitar o que é aceito, como por exemplo um anúncio de bebidas alcoólicas não é aceitável para menores de 18 anos.

O processo de mapear as entradas em relação aos termos linguísticos é feito por meio de

funções de pertinência definidas por especialistas. A partir dela, regras do raciocínio nebuloso são elaboradas, específicas por anúncio, para serem usadas na inferência da relevância do anúncio com base nos valores de entrada.

Define-se então o seguinte algoritmo:

**ALGORITMO PARA O MÉTODO DE SELEÇÃO DE ANÚNCIOS PELA SIMILARIDADE DOS INTERESSES:**

1. Dado o público-alvo do estabelecimento
2. Para cada anuncio:
  - a) Se o gênero do público-alvo do anúncio for diferente de unissex:
    - i) Verificar se o gênero do público-alvo do estabelecimento é o mesmo do público-alvo do anúncio, caso contrário exclui o anúncio da lista e passa para o seguinte.
    - b) Definir a proximidade entre o público-alvo do estabelecimento com o público-alvo do anúncio.
3. Ordenar os anúncios de forma crescente em relação à maior proximidade entre os públicos-alvo.
4. Verifica se o anúncio com maior proximidade já foi exibido.
  - a) Se a condição for verdadeira, repete o procedimento para o anúncio seguinte.
  - b) Se a condição não for verdadeira, seleciona o anúncio para exibição.

## 4.5 Híbridização dos métodos de seleção de anúncios

Para minimizar as limitações dos métodos anteriores propõe-se a combinação de seus resultados em um processo conhecido por híbridização de algoritmos [5]. O objetivo desse processo é minimizar as limitações que esses algoritmos teriam quando usados em separado e ao mesmo tempo aproveitar suas vantagens individuais [32].

Um dos problemas do método de recomendação pela similaridade dos interesses é a sua incapacidade em diferenciar dois produtos similares, mas direcionados a pessoas diferentes. Considere o exemplo de dois anúncios de uísque: um deles é o uísque *Johnnie Walker Blue Label* e o outro o uísque *Old Eight*. Ambos podem ter suas categorias de interesses representadas por:

$$I = \{\text{bebidas alcoólicas, festas}\}$$

O cálculo do índice de *Jaccard* para ambos retornaria o mesmo valor, porém diferentes percepções de relevância podem ser produzidas dependendo da audiência que seja exposta ao anúncio. Apesar das semelhanças por serem um mesmo produto, há uma distinção entre o público-alvo desses produtos. O *Johnnie Walker* é direcionado a pessoas de classe social relativamente alta, enquanto o *Old Eight* é um produto mais popular. O método de recomendação pela similaridade não é capaz de fazer essa distinção.

A diferenciação entre produtos de um mesmo tipo pode ser feita de diferentes formas, uma delas é considerar as características únicas de cada produto. São elas quem o diferencia de outros produtos em seu segmento. No exemplo dos uísques, poderia se considerar o malte utilizado, o tempo de engarrafamento, o tipo de água utilizada, dentre outros. Entretanto, definir essas informações para cada produto seria um processo complexo e de difícil escalabilidade. Utilizar o público-alvo, por ser uma informação comum a todos os produtos, torna essa tarefa viável. Aplicando ao exemplo anterior, tem-se que é possível fazer a diferenciação entre os produtos com base na capacidade financeira do consumidor, ou seja, um anúncio de um uísque com valor aproximado de R\$ 800,00 dificilmente se aplicará a um consumidor de baixa renda.

Por outro lado, apenas a informação sobre o público-alvo de um produto pode não ser suficiente para identificar anúncios relevantes. Por exemplo, considere um grupo de pessoas presente em um estabelecimento com público-alvo de pessoas entre 25 a 40 anos, do sexo masculino e de classe econômica A1 e A2. O anúncio de uma caminhonete direcionado a homens de 30 a 40 anos e de classe econômica A1, seria considerado relevante? Para o algoritmo de classificação essas pessoas pertencem ao público-alvo do produto, porém a resposta para essa pergunta depende de diferentes aspectos. O estilo de carro que essas pessoas preferem é um deles, pois elas podem se interessar apenas por carros esportivos, por exemplo. O uso de categorias de interesse pode ajudar a minimizar esse problema.

Outro problema minimizado pela hibridização dos métodos é que o método de recomendação pela similaridade tem suas recomendações limitadas aos anúncios com categorias relacionadas aos interesses do estabelecimento, limitando a quantidade de anúncios que podem ser recomendados. Através da hibridização, diminui-se a dependência em relação a essas categorias de interesses, pois novas informações serão agregadas na recomendação.

Neste trabalho os métodos serão combinados utilizando uma técnica conhecida como

combinação de recursos ou *feature combination* [5]. Nessa técnica os resultados de cada método são somados para produzir a recomendação final. Um cenário de utilização do método é ilustrado na Figura 4.2.

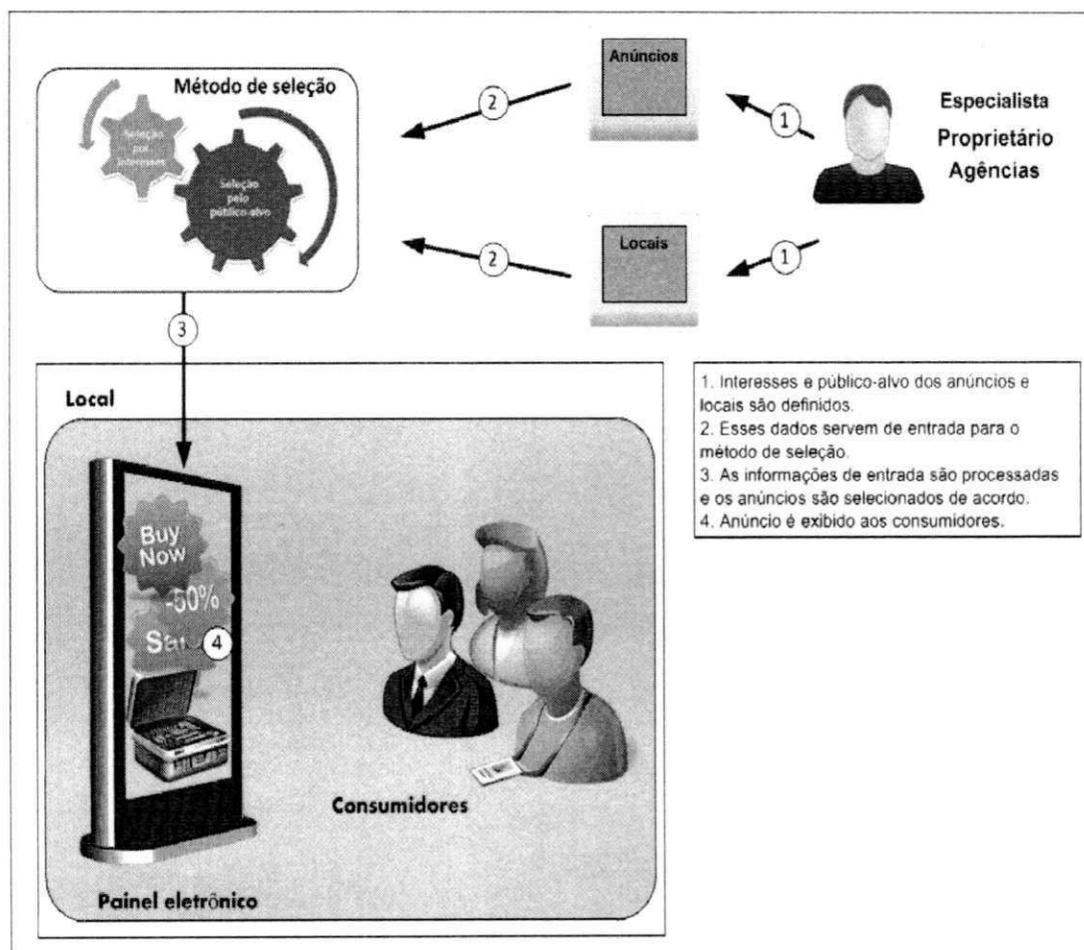


Figura 4.2: Cenário de utilização do modelo.

Diante ao que foi exposto, propõe-se o seguinte algoritmo:

#### ALGORITMO PARA O MÉTODO DE SELEÇÃO DE ANÚNCIOS PELA SIMILARIDADE DOS INTERESSES:

1. Dado os interesses e o público-alvo relacionado ao estabelecimento
2. Para cada anúncio:
  - a) Calcular a similaridade entre os interesses do anúncio e do estabelecimento.
  - b) Calcular a proximidade entre o público-alvo do anúncio e do estabelecimento.

- c) Agregar os valores obtidos.
3. Ordenar os anúncios de forma crescente em relação ao valor agregado.
4. Verificar se o anúncio com maior valor agregado já foi exibido.
  - a) Se a condição for verdadeira, repete o procedimento para o anúncio seguinte.
  - b) Se a condição não for verdadeira, seleciona o anúncio para exibição.

## 4.6 Considerações sobre os métodos

Os métodos propostos têm suas recomendações ou suas definições influenciadas por alguns aspectos. Nessa seção expõem-se algumas deficiências identificadas e como elas afetam o comportamento dos mesmos.

O método de similaridade apresenta dois problemas que afetam diretamente suas recomendações. O primeiro, denominado superespecialização, impede que anúncios que não tenham alguma categoria de interesse relacionada às categorias de interesse do estabelecimento sejam recomendados [44]. Isso minimiza a capacidade desse método em recomendar anúncios diversificados. Outro problema desse método é que os interesses da audiência podem não ser suficientes para diferenciar suas reais preferências. Por exemplo, o interesse de uma pessoa por relógios não será suficiente para identificar se ele prefere relógios Rolex ou Cassio.

Para o método de recomendação pela proximidade entre os públicos-alvo, supõe-se que o uso de apenas essa informação não é suficiente para identificar as preferências dos consumidores. Como descrito anteriormente, o público-alvo é uma informação que facilita a diferenciação entre dois produtos de um mesmo tipo, mas de diferentes segmentos, enquanto os interesses fornecem mais detalhes sobre as preferências do consumidor.

Um problema comum a todos os métodos é a dependência do uso de conhecimento especialista para classificar os interesses associados aos estabelecimentos e anúncios, e para definir as funções de pertinência e regras de raciocínio nebuloso. No entanto, como a quantidade de anúncios e estabelecimentos que precisam ser definidos é relativamente pequena e praticamente não requer alterações no decorrer de seu uso, acredita-se que essa não será uma limitação grave. Além disso, essas informações são de conhecimento de diferentes pessoas, como proprietários dos estabelecimentos, empresas anunciantes, profissionais de

---

publicidade, facilitando assim sua obtenção.

# Capítulo 5

## Validação empírica

### 5.1 Introdução

Para analisar a efetividade dos métodos propostos é preciso avaliar como os consumidores reagem aos anúncios selecionados. Uma das formas de realizar essa avaliação é requisitar a esses consumidores que indiquem quanto esses anúncios foram de interesse [13].

Para realizar esse processo foi conduzido um experimento onde os métodos propostos foram testados e seus resultados comparados aos de outros métodos de propósito similares encontrados na literatura ou em uso em sistemas comerciais de sinalização digital. Para uma maior acurácia nos resultados, o experimento foi realizado em duas etapas. A primeira etapa teve como objetivo identificar possíveis vieses e erros que poderiam afetar a condução do experimento. A segunda etapa, que consiste num aprimoramento da primeira execução do experimento, teve como objetivos específicos:

- Analisar a efetividade das recomendações de anúncios feitas pelos métodos propostos;
- Validar os métodos propostos, comparando seus resultados aos de outros métodos encontrados na literatura;
- Avaliar a viabilidade em usar a informação sobre o público-alvo e interesses de um estabelecimento, descritas por especialistas, para inferir as características de seus consumidores.

## 5.2 Seleção das variáveis

Com base no objetivo do experimento, as seguintes variáveis são definidas para este experimento:

**Variável independente:**

1. Método de seleção dos anúncios: abordagens que fazem a escolha do anúncio a ser exibido ao entrevistado.

**Variável dependente:**

1. Nível de interesse: variável que representa o nível de interesse do anúncio exibido para o entrevistado, segundo sua avaliação.

## 5.3 Condução do experimento

Houve a participação de 112 pessoas no experimento. Dos entrevistados, 65 eram homens e 47 mulheres, com a menor idade de 16 e a maior de 56 anos. Com essa amostra é possível obter dados com um nível de confiança de 95% dentro de um intervalo de confiança de 10% [18]. Como o cenário da sinalização digital não é excludente em relação as pessoas expostas ao anúncio, não foram feitas restrições às características que os participantes do experimento deveriam possuir.

Os entrevistados foram expostos à exibição de diversos anúncios e requisitados a avaliá-los, utilizando uma escala qualitativa para indicar quanto o anúncio havia sido de interesse. O experimento foi conduzido por duas semanas entre as datas de 7 de setembro à 21 de setembro de 2011. Foram registradas 2688 avaliações em 84 dos 112 anúncios disponíveis.

A realização do experimento se deu através da Internet, pois alguns fatores limitaram sua execução *in loco*. Como o cenário do estudo considera locais equipados com displays eletrônicos, para realizar a avaliação *in loco* seria preciso reunir grupos de pessoas e deslocá-las a diferentes locais. Isso impõe dificuldades de logística e horário. Considerando que o viés em avaliar os entrevistados individualmente é mínimo, se comparado à avaliação presencial e em grupo, a realização do experimento pela Internet foi a alternativa mais viável.

Como o trabalho considera que um anúncio é exibido para múltiplas pessoas presentes em um mesmo local e os entrevistados foram avaliados individualmente e em locais distintos, o comportamento de grupo foi simulado. Para isso, os entrevistados foram requisitados a informar quais estabelecimentos eles haviam visitado no ano, com essa informação foram formados grupos com as pessoas que indicaram ter visitado um mesmo local. Dessa forma, simulando a presença dos consumidores em um mesmo local. A esse processo foram definidas algumas restrições:

- Cada grupo tinha relação com apenas um local;
- A alocação dos entrevistados aos grupos foi feita de forma aleatória à medida que eles participaram do experimento, esse é o princípio base para blocagem em experimentos [12];
- Quando não houve a possibilidade do entrevistado ser alocado em um grupo, um novo foi criado, caso ainda não existisse;
- Os grupos foram limitados a cinco integrantes, exceto em situações onde não houve a possibilidade de criar um novo grupo. Isso assegura que haverá uma quantidade diversificada de locais;
- Todos os integrantes de um grupo viram o mesmo anúncio.

### **5.3.1 Classificação dos anúncios e estabelecimentos**

Foi utilizada uma base com 112 anúncios relacionados a produtos reais e dos mais variados assuntos. Dois especialistas em publicidade foram solicitados a definir as categorias de interesse e classificar os anúncios e estabelecimentos em relação a elas. Foram definidas 33 categorias de interesse a partir de informações obtidas em sites de premiações de anúncios e no Conar (Conselho de auto-regulamentação Publicitária)<sup>1</sup>.

Os especialistas também classificaram o público-alvo dos anúncios e estabelecimentos. Em ambos os processos foi priorizada a informação que os especialistas encontraram em fontes especializadas, como sites e revistas. Na sua ausência foi utilizado o conhecimento

---

<sup>1</sup><http://www.conar.org.br/>

técnico desses profissionais. Para minimizar o viés em haver muitos anúncios direcionados a apenas um determinado público-alvo, os publicitários foram orientados a selecionar anúncios para diferentes públicos-alvo.

### 5.3.2 Métrica para avaliação dos métodos

Analisando o estado da arte percebe-se que a maioria dos métodos identificados são avaliados em comparação à seleção aleatória, formato mais popular de seleção de conteúdo em sinalização digital. Percebe-se também que uma das métricas mais usadas para avaliá-los é em relação à precisão dos resultados.

Neste trabalho foi utilizado um formato de avaliação semelhante, incluindo também a comparação a dois algoritmos de filtragem colaborativa. Essa classe de algoritmos também é utilizada em trabalhos acadêmicos que visam melhorar a seleção de anúncios, além disso, os métodos propostos nesta dissertação podem ser classificados como algoritmos de recomendação. Tentou-se o acesso a algoritmos propostos na área acadêmica no cenário de sinalização digital, mas sem sucesso.

Diante disto, nesse experimento foram testados seis métodos:

1. Três propostos nesse estudo: método de seleção pela similaridade entre os interesses dos anúncios e estabelecimentos, método baseado na proximidade entre o público-alvo dos anúncios e estabelecimentos e a combinação de ambos (Capítulo 4);
2. Duas abordagens de filtragem colaborativa: filtragem colaborativa por correlação de *Pearson* e filtragem colaborativa por *Slope One*;
3. Seleção aleatória.

Cada método selecionou para exibição três anúncios por entrevistado, que o avaliou em uma escala qualitativa com os valores: desinteressante, pouco interessante e interessante. Uma vez o anúncio avaliado, caso um método o selecionasse novamente, ele não seria exibido e sua avaliação seria aproveitada. Espera-se com isso minimizar o viés em exibir anúncios repetidos, que poderia vir a torná-lo ofensivo, no sentido de ser irritante, e reduzir o interesse do consumidor por ele [6].

Para comparar os métodos, foram analisadas as avaliações que os entrevistados fizeram nos anúncios selecionados. Os resultados foram medidos utilizando duas métricas para avaliar a qualidade das recomendações: precisão e taxa de erro [15], [13].

A representação matemática para precisão e taxa de erro é a seguinte:

$$P = \frac{pv}{t}$$

$$TE = \frac{fp}{t}$$

Onde:

- P = Precisão;
- TE = Taxa de erro;
- pv = Positivo verdadeiro, onde o anúncio exibido é avaliado como relevante;
- fp = Falso positivo, onde o anúncio exibido é avaliado como irrelevante;
- t = Total de recomendações.

Precisão é a fração de documentos selecionados para exibição e que foram relevantes ao usuário. Enquanto a taxa de erro refere-se à fração de documentos selecionados, mas que foram irrelevantes ao usuário.

## 5.4 Primeira etapa

A primeira etapa do experimento teve como objetivo identificar possíveis problemas e vieses em sua condução e que poderiam influenciar os resultados finais da pesquisa. Nessa etapa houve a participação de 14 pessoas, divididos igualmente entre 7 homens e 7 mulheres. Após a execução do experimento os entrevistados foram contatados para expor sua opinião em relação ao estudo. A partir desse contato foram identificados os seguintes problemas:

- Foi exibida uma grande quantidade de anúncios, o que tornou o experimento entediante;

- Houve problemas em entender o significado da palavra relevante, usada na escala de avaliação dos anúncios, e os objetivos do experimento;
- O uso de uma escala de quatro valores para avaliar os anúncios ocasionou problemas para interpretar as diferenças entre esses níveis.

Inicialmente foi definido que cada método faria a seleção de 10 anúncios e ao fim seriam avaliados aproximadamente 60 anúncios por entrevistado. Essa grande quantidade de anúncios ocasionou uma frustração aos entrevistados, que a partir do momento que consideraram o estudo entediante passaram a avaliar os anúncios de qualquer maneira. O viés causado prejudicou principalmente os últimos métodos usados na seleção dos anúncios. Para minimizar esse problema foi reduzida a quantidade de anúncios selecionados de 10 para 3 e o métodos passaram a se alternar na seleção dos anúncios, ao contrário do antigo modelo sequencial utilizado.

Outra reclamação dos entrevistados foi em relação ao uso da palavra relevância na escala de avaliação dos anúncios. Para eles, não ficou clara a semântica da palavra e o que estava sendo avaliado. O uso de quatro escalas para avaliação também foi questionado, pois não ficou entendida a diferença entre os níveis: irrelevante, pouco relevante, relevante e muito relevante. Para diminuir as dúvidas a palavra relevante foi trocada por interessante, por sugestão dos entrevistados, e houve uma melhora no texto de explicação do experimento. A escala de avaliação também foi alterada de 5 para 3 valores: desinteressante, pouco interessante e interessante, pois considera-se que há uma menor ambiguidade nessa escala.

## 5.5 Segunda etapa

Após as correções dos problemas identificados na primeira etapa teve início a segunda etapa do experimento. Foram entrevistadas 112 pessoas que passaram pelas seguintes atividades:

1. Cada entrevistado foi solicitado a informar diante de uma lista de estabelecimentos, quais eles visitaram no último ano. Essa informação foi usada para alocar o entrevistado em um grupo relacionado a algum dos locais informados e para inferir informações sobre o entrevistado;

2. Com a informação sobre os locais visitados, o sistema inferiu os dados demográficos e interesses do entrevistado a partir dos dados de marketing desses estabelecimentos. Foi solicitado que as informações inferidas erroneamente fossem corrigidas;
3. Por fim, os anúncios foram selecionados pelos métodos testados. Para cada anúncio visto o entrevistado o avaliava indicando o quanto o anúncio havia sido de interesse e o sistema armazenava essa informação em uma base de dados. A informação armazenada incluía: o método que selecionou o anúncio, o anúncio exibido, o entrevistado que o avaliou e a nota de avaliação.

## 5.6 Resultados

### 5.6.1 Inferência do perfil do entrevistado a partir do público-alvo e categorias de interesse do estabelecimento

Essa avaliação foi utilizada para verificar se os consumidores de determinado estabelecimento possuíam características similares às definidas pelos publicitários para o público-alvo desses locais. Os resultados dessa avaliação indicarão a viabilidade em inferir o perfil dos entrevistados a partir dos dados que descrevem o estabelecimento onde elas se encontram.

#### Inferência dos interesses

Buscou-se verificar se as pessoas que indicaram ter visitado determinado estabelecimento, interessavam-se pelas categorias de interesse associadas ao local. Para isso foi medida a diferença entre os interesses inferidos pelo sistema e os interesses após correção pelo entrevistado.

Para medir a diferença entre os dois, foi usada a distância de *Hamming* [20] que indica a quantidade de modificações que o entrevistado fez nos interesses inferidos para se chegar aos interesses corretos. Os resultados são apresentados na Figura 5.1.

Pelo gráfico observa-se que para 26.7% dos entrevistados seus interesses foram inferidos corretamente. O erro médio foi de 3.33 modificações, seja acrescentando ou retirando categorias de interesse, e no pior caso houve 14 alterações, havendo apenas uma ocorrência. Considerando os casos onde foram necessárias de 1 a 4 modificações, o percentual para essas

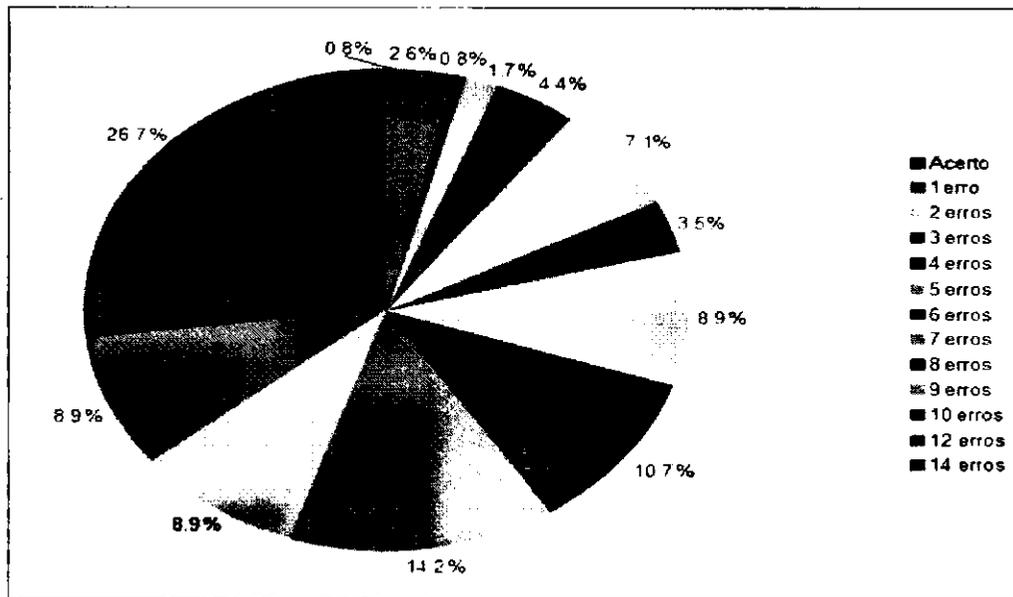


Figura 5.1: Percentuais para o acerto e quantidade de erros para a inferência dos interesses do entrevistado.

ocorrências foram de 42,7%, valores aceitáveis para esse trabalho em virtude da definição do público-alvo aceitar intervalos de valores para classificar os valores aceitáveis.

### Inferência da idade

Outra informação inferida do entrevistado foi a sua idade. Pela Figura 5.2, percebe-se que para 16% dos entrevistados esse valor foi inferido corretamente. Apesar do pequeno percentual, ao analisar os casos onde a inferência falhou, percebe-se que para 70,4% dos entrevistados a inferência errou no máximo 5 anos. Como o público-alvo em sua definição aceita margens, considera-se esse um valor aceitável. As piores variações de idade foram de 27 anos, ocorrida com um entrevistado, seguida por 14 anos que ocorreu a dois entrevistados.

### Inferência do gênero

O gênero dos entrevistados também foi inferido, havendo a possibilidade do sistema inferir a informação corretamente, com erro ou haver falha. Como havia estabelecimentos sem diferenciação de gênero, direcionados a um público unissex, nessas situações o sistema não iria conseguir inferir corretamente o sexo do entrevistado.

Os resultados apresentados na Figura 5.3 mostram que para 94,6% dos entrevistados o

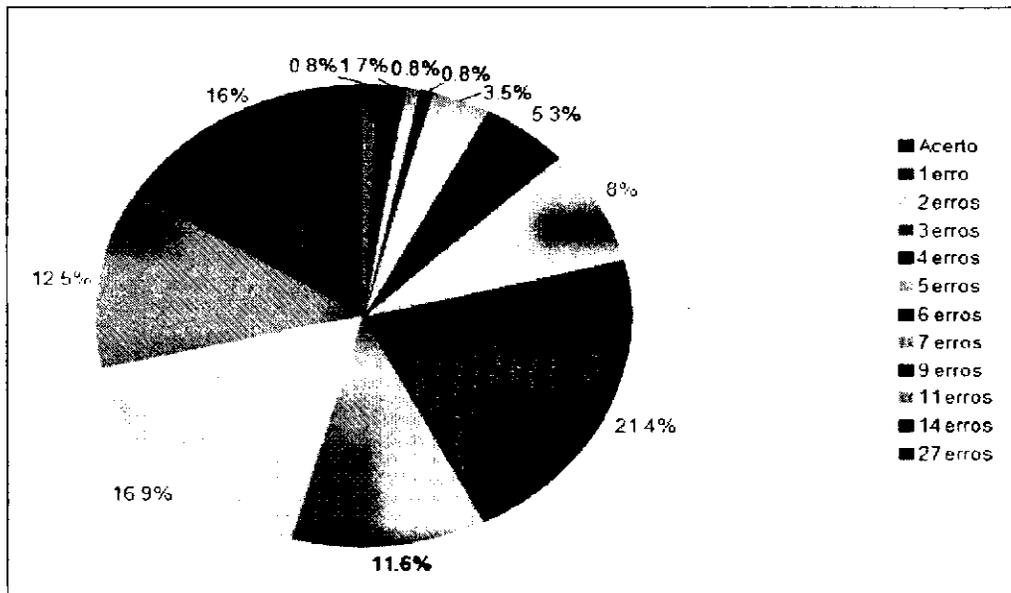


Figura 5.2: Percentuais para o acerto e a diferença em anos entre a idade inferida e a idade real do entrevistado.

sexo foi inferido corretamente e para 5.4% dos entrevistados a inferência foi incorreta, não havendo nenhuma ocorrência para sexo não identificado.

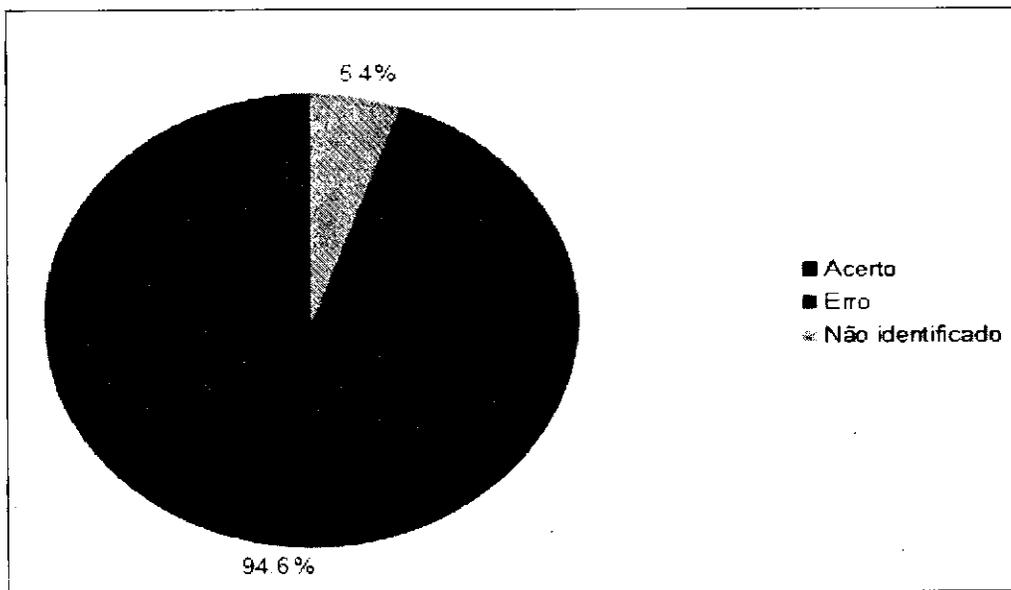


Figura 5.3: Percentuais de acerto e erro para a inferência do genero.

### Inferência da classe social

A inferência da classe social indica a qual classe econômica o entrevistado possivelmente pertence. Os valores aceitos foram baseados no critério de classificação econômica Brasil [1].

Pelos resultados apresentados na Figura 5.4, percebe-se que para 51.3% dos entrevistados sua classe econômica foi inferida corretamente. Para 33% houve um erro de uma classe social para cima ou para baixo e para 11.6% o erro foi de duas classes sociais. O pior dos casos foi a variação de três classes sociais, para mais ou menos, que ocorreu com 3.5% dos entrevistados.

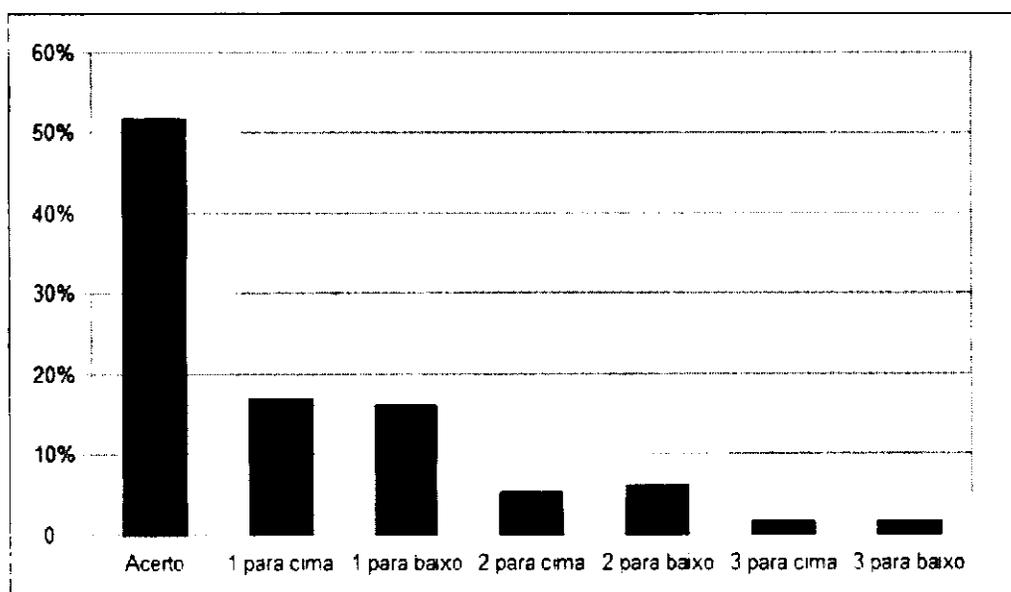


Figura 5.4: Percentuais de acerto e erro para a inferência do gênero.

## 5.6.2 Comparação entre os métodos

### Resultados da precisão e taxa de erro

Para comparar os métodos testados foram analisadas as métricas de precisão e taxa de erro referentes às avaliações que os entrevistados realizaram nos anúncios exibidos. Os resultados são apresentados nas tabelas 5.1 e 5.2.

A precisão indica o percentual de anúncios exibidos e avaliados como interessantes pelos

Aleatório	Método de similaridade	Método de proximidade do P.A.	Método híbrido	F.C. por correlação de Pearson	F.C por Slope One
38%	50%	33.9%	48.8%	49.4%	54.1%

Tabela 5.1: Precisão das abordagens.

Aleatório	Método de similaridade	Método de proximidade do P.A.	Método híbrido	F.C. por correlação de Pearson	F.C por Slope One
35.1%	28.2%	36.3%	24.1%	27.9%	19.3%

Tabela 5.2: Taxa de erro das abordagens.

entrevistados. Pelos resultados percebe-se que o algoritmo *Slope One* obteve o melhor resultado, com 182 dos 336 anúncios selecionados avaliados dessa forma. O método de seleção pela similaridade entre os interesses dos anúncios e estabelecimentos obteve o segundo melhor resultado, com 168 anúncios exibidos, seguido pelo algoritmo de filtragem colaborativa por correlação de *Pearson* com 166 anúncios, método híbrido com 164 anúncios, seleção aleatória com 128 anúncios e o método baseado na proximidade entre os públicos-alvos com 114 anúncios.

Analisando a taxa de erro, novamente o algoritmo *Slope One* obteve o melhor resultado ao recomendar a menor quantidade de anúncios avaliados como desinteressantes, selecionando para exibição 65 anúncios que foram avaliados com essa escala. No entanto, observa-se uma melhora no desempenho do método híbrido que selecionou menos anúncios desinteressantes em comparação ao método por similaridade e ao algoritmo de filtragem colaborativa por correlação de *Pearson*. O método híbrido selecionou 81 anúncios, enquanto o método por similaridade selecionou 95 anúncios, seguido pela correlação de *Pearson* com 94 anúncios. Sem modificações nas últimas posições, a seleção aleatória recomendou 118 anúncios e o método baseado na proximidade entre os públicos-alvos 122 anúncios desinteressantes.

Em relação aos anúncios avaliados como pouco interessantes, apesar de não entrarem no cálculo, essas avaliações indicam anúncios que são indiferentes aos consumidores, mas se mostram melhores que os anúncios avaliados como desinteressantes. Dos seis métodos testados apenas dois tiveram a quantidade de anúncios avaliados como pouco interessantes

maiores que as seleções de anúncios avaliados como desinteressantes: o método híbrido e *Slope One*. Esse resultado indica uma maior eficácia desses métodos em filtrar anúncios irrelevantes aos consumidores. Pelo gráfico da Figura 5.5 também é possível analisar os resultados da precisão e taxa de erro.

Apesar do resultado inferior do método híbrido em comparação ao algoritmo *Slope One*, esse método se mostra como uma alternativa viável principalmente por ter conseguido resultados próximos aos de *Slope One*, sem necessitar de interação explícita com o consumidor. Como exposto na Introdução desta dissertação, em um cenário real de sinalização digital esse tipo de interação é bastante difícil.

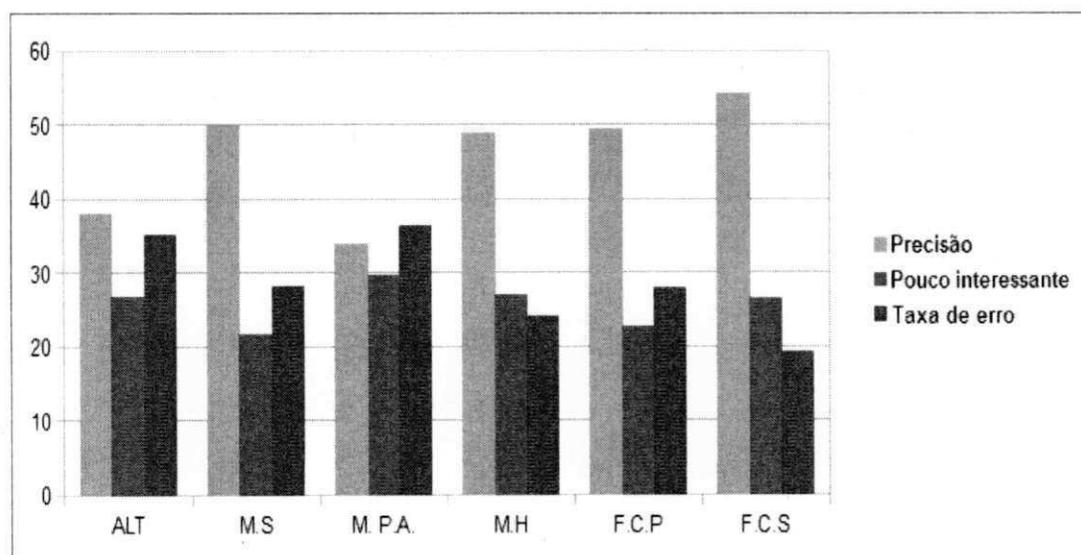


Figura 5.5: Gráfico com os percentuais da precisão, taxa de erro e dos anúncios avaliados como pouco interessantes. Legenda: alt - Aleatorio, M.S - Método de seleção pela similaridade dos interesses, M. P.A. - Método de seleção pela proximidade entre os públicos-alvos, M.H - Método híbrido, F.C. P. - Filtragem colaborativa por correlação de *Pearson*, F.C. S. - Filtragem colaborativa por *Slope One*.

## 5.7 Análise e interpretação dos resultados

### 5.7.1 Inferência do perfil do entrevistado

Os resultados da inferência do perfil do entrevistado indicam que há uma relação entre os dados que descrevem o público-alvo dos estabelecimentos e seus consumidores. Foram obti-

dos percentuais de acerto consideráveis, e mesmo em situações onde esses percentuais foram abaixo do esperado, a variação ocorrida foi pequena na maioria dos casos e considerada aceitável para o cenário deste estudo.

Apesar dos bons resultados é importante compreender as situações onde esse modelo não é adequado, podendo assim explicar os valores atípicos que ocorreram no experimento. Analisando os resultados e por levantamentos na literatura considera-se que o processo de inferência pode errar nas seguintes situações:

1. Uma pessoa estar presente em um estabelecimento por estar acompanhando outra pessoa que de fato possui interesse pelo local;
2. Estar presente no local, mas exercer alguma atividade nele, como por exemplo, um trabalho. Nessa situação pode ser que essa pessoa não tenha interesse pelo estabelecimento;
3. Visitar um local para a compra de um presente a outra pessoa;
4. O estabelecimento foi visitado, mas não era de interesse do consumidor.

Além dos casos acima, para cada um dos dados inferidos há situações específicas que podem explicar os cenários mais adequados ao seu uso ou o porquê de algumas inferências terem sido erradas. Essas situações serão analisadas a seguir.

### **5.7.2 Análise da inferência do gênero**

Para 94.6% dos entrevistados seu gênero foi inferido corretamente. Esse resultado evidencia a forte relação que há entre o que é oferecido nos estabelecimentos e o gênero de quem o consome.

Analisando os resultados, foi identificado que a eficácia dessa abordagem é mais evidente em locais onde o público-alvo é formado predominantemente por pessoas de um mesmo sexo, como exemplo em lojas de roupa específicas por sexo. Em locais onde há uma menor diferenciação, como restaurantes, bares e supermercados, a inferência é mais susceptível a erros. O fato de não haver nenhuma ocorrência de gênero não identificado mostra que apesar de haverem locais destinados a ambos os sexos, todos os entrevistados visitaram no mínimo um estabelecimento específico para um gênero.

### 5.7.3 Análise da inferência da idade

A inferência da idade obteve um percentual de acerto de 16%. O resultado inicialmente parece ser ruim, mas analisando os erros, percebe-se que para 70.5% dos entrevistados o erro máximo foi de apenas 5 anos, ou seja, para esses entrevistados a variação entre a idade inferida e a idade correta foi de 1 a 5 anos. Para o cenário do trabalho, isto é considerado aceitável.

No entanto, analisando os resultados, essa inferência não se mostrou muito eficaz para entrevistados com idade inferior a 18 e superior a 45 anos, onde houve os maiores desvios. Para compreender o porquê da inferência não ter se mostrado adequada a essas pessoas, foi observado na descrição do público-alvo dos estabelecimentos que apenas uma pequena quantidade de locais foram definidos com público-alvo direcionado a pessoas dessas idades.

### 5.7.4 Análise da inferência da classe econômica

Os resultados da inferência da classe econômica são apresentados na Figura 5.4 onde conclui-se que para 51.7% dos entrevistados a classe econômica foi inferida corretamente. Analisando os erros, a variação de uma classe econômica para cima ou para baixo ocorreu a 33% dos entrevistados. O bom resultado é um indício que demonstra a relação entre a classe econômica dos consumidores e os produtos que eles consomem.

Há a possibilidade de haverem valores atípicos, como uma pessoa de uma classe econômica inferior ou superior ao público-alvo de um estabelecimento, frequentá-lo. No entanto, além das situações descritas anteriormente, onde a inferência pode não ser adequada, se essa pessoa está presente no estabelecimento isso é um indicativo de que ela se interessa por algo relacionado a ele.

### 5.7.5 Avaliação dos métodos

Pela análise dos resultados da precisão (Tabela 5.1) e taxa de erro (Tabela 5.2), percebe-se a superioridade do algoritmo *Slope One* em comparação as outras abordagens. Apesar do resultado inferior, o método híbrido obteve um desempenho satisfatório, mantendo-se próximo aos resultados do algoritmo.

Analisando a métrica de precisão percebe-se uma diferença insignificante entre os resultados do método de similaridade, algoritmo de filtragem colaborativa por correlação de *Pearson* e o método híbrido. A seleção aleatória obteve a quinta colocação, seguida pelo método de proximidade dos públicos-alvo que obteve o pior resultado.

A análise da taxa de erro demonstra que, excluindo o algoritmo *Slope One*, o método híbrido se mostrou mais eficaz em filtrar anúncios irrelevantes que as outras abordagens. Em comparação à seleção aleatória houve uma diferença significativa de 10 pontos percentuais entre os resultados. Novamente o método de proximidade dos públicos-alvo obteve o pior resultado, sendo o único com a taxa de erro superior à precisão.

Para um entendimento correto dos resultados é fundamental analisar conjuntamente a precisão e a taxa de erro. Por exemplo, ao analisar apenas os resultados da precisão, percebe-se que o método de seleção pela similaridade dos interesses selecionou mais anúncios avaliados como interessantes do que o método híbrido (quatro a mais). No entanto, verificando a taxa de erro muda-se a interpretação dos resultados. Nessa métrica, o método híbrido selecionou menos anúncios irrelevantes do que o método de similaridade (dezoito a menos). Analisando ambos, conclui-se que apesar do método híbrido ter tido resultado um pouco inferior para a precisão, analisando a taxa de erro a variação foi significativa e favorável ao método híbrido, o que lhe garante um desempenho mais estável (Figura 5.5).

## 5.8 Conclusão dos resultados

Os bons resultados obtidos com a inferência do perfil dos entrevistados demonstram que: *i)* os dados que descrevem o público-alvo e as categorias de interesse relacionadas a um estabelecimento condizem com as características de seus consumidores; *ii)* em razão do que foi identificado no ponto *i)* percebe-se que há uma homogeneidade das características dos consumidores de um estabelecimento. O uso dessa informação pode ser uma forma de minimizar a incerteza sobre os dados dos consumidores, principalmente em situações onde a obtenção de informações explícitas dos usuários pode não ser viável ou desejável.

Com os resultados das métricas de precisão e taxa de erro do método híbrido, e sua posterior validação em comparação a métodos de propósitos similares identificados no estado da arte, percebe-se a viabilidade dessa abordagem na seleção de anúncios. Esse método se

mostrou mais efetivo na filtragem de anúncios irrelevantes e na seleção de anúncios relevantes obteve um desempenho praticamente igual ao método de seleção pela similaridade e ao algoritmo de filtragem colaborativa por correlação de Pearson. A seleção aleatória se mostrou pouco efetiva na seleção de anúncios relevantes e filtragem dos anúncios irrelevantes, mantendo-se a frente somente do método de seleção pela proximidade dos públicos-alvos.

Pelos resultados acredita-se que o uso de apenas os dados do público-alvo para recomendar anúncios não é viável. No entanto, quando combinado ao uso dos interesses, percebe-se que o método híbrido possibilitou uma diminuição na taxa de erro em comparação aos resultados individuais dos algoritmos. Pode-se atribuir isso ao fato do método híbrido encontrar anúncios de interesse do usuário e direcionados ao seu público-alvo mais efetivamente que as outras abordagens.

## Capítulo 6

### Conclusão

Nesta dissertação foi proposto um método para melhorar a seleção de anúncios em sinalização digital. Esse método é resultado da agregação de dois conceitos usados pela área de marketing no processo de personalização de produtos: os interesses dos consumidores e o público-alvo dos produtos. Seu uso possibilita encontrar anúncios relacionados a assuntos de interesse dos consumidores, mas que também sejam relevantes a suas condições financeiras, idade e gênero.

Para obter os dados dos consumidores necessários ao método, optou-se pelo uso de dados do público-alvo do estabelecimento onde essas pessoas estão presentes. Essa é uma informação acessível à sinalização digital e seu uso minimiza a necessidade de interação com os consumidores e a dependência de dispositivos tecnológicos, problemas comuns às abordagens similares encontradas na literatura. Posteriormente o experimento realizado comprovou que os dados do público alvo dos locais representam, com boa acurácia, as características reais das pessoas que frequentam esse local.

O método proposto foi avaliado em um experimento com a participação de 112 entrevistados. Essas pessoas foram requisitadas a avaliar as recomendações produzidas pelo método proposto e também por outras abordagens identificadas na literatura ou usadas em sistemas comerciais de sinalização digital. Os resultados mostram que o método proposto foi efetivo no processo de recomendação de anúncios, principalmente ao reduzir a exibição de anúncios com potencial irrelevância. Ficou comprovada também a superioridade do método em comparação a seleção aleatória. Em comparação com a filtragem colaborativa, os resultados apesar de inferiores foram próximos, e em alguns casos superiores, mas com a vantagem de

não depender de entradas explícitas da audiência.

Propõe-se como trabalho futuro utilizar fontes de informações mais confiáveis e automatizadas para definir os dados que descrevem os anúncios e estabelecimentos usados pelo método. Espera-se com isso minimizar a necessidade de conhecimento especialista. O uso de dados de redes sociais que indicam as características dos frequentadores de um estabelecimento, como a área do comerciante no *Foursquare*, ou o contato com as agências de publicidade que definem os anúncios, podem ser usadas para obter informações mais detalhadas e precisas sobre o público-alvo desses produtos. Propõe-se também agregar informações de contexto, como a localização geográfica, horário e a temperatura, com o objetivo de tornar as recomendações mais efetivas.

Espera-se no futuro validar o trabalho *in loco*, utilizando painéis eletrônicos instalados em estabelecimentos reais e com anunciantes também reais. Dessa forma o método proposto poderá ser avaliado em condições do cotidiano, sob as influências do ambiente do estabelecimento. Além disso, também poderão ser avaliados aspectos relacionados a escalabilidade da proposta, bem como sua usabilidade para os consumidores.

O trabalho contribui para a área de publicidade pervasiva ao propor um método para recomendar anúncios mais efetivos que usa informações apoiadas em conceitos de marketing bem definidos. Não foram encontrados trabalhos na área que utilizassem metodologia semelhante. Além disso, o método proposto reduziu a necessidade de interação com os consumidores e eliminou a dependência de dispositivos tecnológicos, como encontrado em trabalhos semelhantes. Com poucas modificações acredita-se que essas informações poderão ser adaptadas a outros cenários, como a recomendação de restaurantes, viagens, dentre outros.

## Bibliografia

- [1] Associação brasileira de empresas de pesquisa. Critério de classificação econômica Brasil. <http://www.abep.org/novo/FileGenerate.ashx?id=257>, 2011.
- [2] GE Belch and MA Belch. *Advertising and Promotion: An Integrated Marketing Communications Perspective*. McGraw-Hill, 6 th edition, 2004.
- [3] J. Bickers. The state of the digital signage industry survey. Technical report, Digital Signage Today, <http://www.digitalsignagetoday.com/whitepapers/79/The-State-of-the-Digital-Signage-Industry-Survey>, 2010.
- [4] P. Borges. *A model of strategy games based on the paradigm of the Iterated Prisoner's Dilemma employing Fuzzy Sets*. PhD thesis, UFSC-EPS, 1996.
- [5] R. Burke. Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12(4):331–370, 2002.
- [6] K. Chan, L. Li, S. Diehl, and R. Terlutter. Consumers' response to offensive advertising: a cross cultural study. *International Marketing Review*, 24(5):606–628, 2007.
- [7] Marketing Charts. Global dooh revenue expected to rise 17%. <http://www.marketingcharts.com/direct/global-doooh-revenue-expected-to-rise-17-16334/>, 2011.
- [8] J.R.C. da Silveira. Publicidade: uma questão de relevância. *VIII Encontro do Círculo de Estudos Linguísticos do Sul*, 2008.
- [9] A. Di Ferdinando, A. Rosi, R. Lent, A. Manzalini, and F. Zambonelli. Myads: A system for adaptive pervasive advertisements. *Pervasive and Mobile Computing*, 5(5):385–401, 2009.

- [10] J. Exeler, M. Buzeck, and J. Müller. eMir: Digital signs that react to audience emotion. In *Proceedings of the 2nd Workshop on Pervasive Advertising, Lübeck*, volume 154, pages 3904–3910, 2009.
- [11] J. Froehlich, M. Chen, I. Smith, and F. Potter. Voting with your feet: An investigative study of the relationship between place visit behavior and preference. *UbiComp 2006: Ubiquitous Computing*, pages 333–350, 2006.
- [12] A.C. Gil. *Como elaborar projetos de pesquisa*. 4<sup>o</sup> edition.
- [13] A. Gunawardana and G. Shani. A survey of accuracy evaluation metrics of recommendation tasks. *The Journal of Machine Learning Research*, 10:2935–2962, 2009.
- [14] J.V. Harrison and A. Andrusiewicz. Using wireless networks to enhance narrowcasting in public spaces. In *Consumer communications and networking conference*, pages 600–604. IEEE, 2004.
- [15] F. Hernández del Olmo and E. Gaudioso. Evaluation of recommender systems: A new approach. *Expert Systems with Applications*, 35(3):790–804, 2008.
- [16] E.M. Huang and E.D. Mynatt. Semi-public displays for small, co-located groups. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 49–56. ACM, 2003.
- [17] C. Ihlström Eriksson, M. Akesson, and J. Hakeröd. Advertising in ubiquitous media environments. In *Proceedings of the 30th Information Systems Research Seminar in Scandinavia, IRIS 30*. Department of Computer Sciences, University of Tampere, Finland, 2007.
- [18] R. Jain. *The art of computer systems performance analysis*. Wiley-India, 2008.
- [19] M. Karam, T. Payne, and E. David. Evaluating bluscreen: Usability for intelligent pervasive displays. In *Pervasive Computing and Applications, 2007. ICPCA 2007. 2nd International Conference on*, pages 18–23. IEEE, 2007.
- [20] N.K. Kasabov. *Foundations of neural networks, fuzzy systems, and knowledge engineering*. The MIT Press, 1996.

- [21] K. Kelsen. *Unleashing the Power of Digital Signage: Content Strategies for the 5th Screen*. Focal Press, 2010.
- [22] M. Khan. *Consumer Behaviour And Advertising Management*. New Age International, 2007.
- [23] J. McCarthy, T. Costa, and E. Liongosari. Unicast, outcast & groupcast: Three steps toward ubiquitous, peripheral displays. In *Ubicomp 2001: Ubiquitous Computing*, pages 332–345. Springer, 2001.
- [24] M. McDonald and I. Dunbar. *Market segmentation: how to do it, how to profit from it*. Butterworth-Heinemann, 2004.
- [25] J. Müller. Traditional and digital signage. *2nd Workshop on Pervasive Advertising*, 2009.
- [26] J. Müller, F. Alt, and D. Michelis. *Pervasive Advertising*. Springer, 2011.
- [27] J. Müller, J. Exeler, M. Buzeck, and A. Krüger. Reflectivesigns: Digital signs that adapt to audience attention. *Pervasive Computing*, pages 17–24, 2009.
- [28] J. Müller and A. Krüger. Mobidic: Context adaptive digital signage with coupons. *Ambient Intelligence*, pages 24–33, 2009.
- [29] C. Narayanaswami, D. Coffman, MC Lee, YS Moon, JH Han, HK Jang, S. McFaddin, YS Paik, JH Kim, JK Lee, et al. Pervasive symbiotic advertising. In *Proceedings of the 9th workshop on Mobile computing systems and applications*, pages 80–85. ACM, 2008.
- [30] K. Philip and A. Gary. *Principles of marketing European edition*, volume 4th.
- [31] K. Phillip and K. Keller. *Marketing Management*. Prentice Hall, 13th edition, 2008.
- [32] M. Pinto. *Sistema de recomendação de produtos em marketing com uso de filtros colaborativos e números fuzzy*. PhD thesis, PUC-RIO, 2011.
- [33] E. Polo. A victoria's secret brasileira. *Isto É - Dinheiro*, 704, 2011.

- [34] D. Quercia, G. Di Lorenzo, F. Calabrese, and C. Ratti. Mobile phones and outdoor advertising: Measurable advertising. *Pervasive Computing, IEEE*, (99):28–36, 2011.
- [35] A. Ranganathan and R.H. Campbell. Advertising in a pervasive computing environment. In *Proceedings of the 2nd international workshop on Mobile commerce*, pages 10–14. ACM, 2002.
- [36] F. Ribeiro, M. Costa, and J. Metrólho. Informing and socializing in interactive spaces: delivering the right information to the right users. In *Proceedings of the 10th WSEAS international conference on Artificial intelligence, knowledge engineering and data bases*, pages 205–210. World Scientific and Engineering Academy and Society (WSEAS), 2011.
- [37] F.R. Ribeiro. *Escalonamento autónomo e sensível ao contexto para ecrãs públicos*. PhD thesis, Universidade do minho - Portugal, 2010.
- [38] F.R. Ribeiro and R. José. Timely and keyword-based dynamic content selection for public displays. In *Complex, Intelligent and Software Intensive Systems (CISIS)*, pages 655–660. IEEE, 2010.
- [39] M. Sharifi, T. Payne, and E. David. Public display advertising based on bluetooth device presence. In *Mobile Interaction with the Real World (MIRW 2006) in conjunction with the 8th International Conference on Human Computer Interaction with Mobile Devices and Services*, 2006.
- [40] D. Sperber and D. Wilson. *Relevance: Communication and cognition*. Harvard University Press, 2nd edition, 1995.
- [41] M. Strohbach, E. Kovacs, and M. Martin. Towards pervasively adapting display networks. In *Proceedings of the 1st Pervasive Advertising Workshop*, 2009.
- [42] M. Strohbach and M. Martin. Towards a platform for pervasive display applications in retail environments. *Pervasive Computing, IEEE*, (99), 2011.
- [43] J. Symonds and M. Khosrow-Pour. *Ubiquitous and Pervasive Computing: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2009.

- 
- [44] R. Torres. *Personalização na Internet: como descobrir os hábitos de consumo de seus clientes, fidelizá-los e aumentar o lucro de seu negócio*. Editora Novatec, 2004.
- [45] M. Weiser. The computer for the 21st century. *Scientific American*, 265(3):94–104, 1991.
- [46] K.M. Yu, C.Y. Yu, B.H. Yeh, C.H. Hsu, and H.N. Hsieh. The design and implementation of a mobile location-aware digital signage system. In *Mobile Ad-hoc and Sensor Networks (MSN)*, pages 235–238. IEEE, 2010.