

**UM MODELO PARA OTIMIZAÇÃO DO RESULTADO DE CAMPANHAS DE
MARKETING DIRETO BASEADO EM ALGORITMOS GENÉTICOS E ANÁLISE
DE RECÊNCIA, FREQUÊNCIA E VALOR**

Eduardo Martins Ribeiro

Universidade Federal do Rio de Janeiro

Instituto de Matemática

Mestrado em Informática

Co-orientador: Antonio Juarez Alencar, D.Phil

Co-orientador: Eber Assis Schmitz, Ph.D

Rio de Janeiro, RJ – Brasil

2005

**UM MODELO PARA OTIMIZAÇÃO DO RESULTADO DE CAMPANHAS DE MARKETING
DIRETO BASEADO EM ALGORITMOS GENÉTICOS E ANÁLISE DE RECÊNCIA,
FREQUÊNCIA E VALOR**

Eduardo Martins Ribeiro

Dissertação submetida ao corpo docente do Instituto de Matemática da
Universidade Federal do Rio de Janeiro – UFRJ, como parte dos requisitos
necessários à obtenção do grau de Mestre em Ciências (M.Sc.) em Informática.

Aprovada por:

Prof. Antonio Juarez Alencar, D.Phil. – Co-orientador

Prof. Eber Assis Schmitz, Ph.D. – Co-orientador

Prof. Fernando S. P. Manso, Ph.D - (UFRJ/NCE)

Prof. Alexandre Marinho, Ph.D. – (UERJ / IPEA)

Rio de Janeiro, RJ – Brasil

2005

FICHA CATALOGRÁFICA

Ribeiro, Eduardo Martins.

Um Modelo Para Otimização do Resultado de Campanhas de Marketing Direto Baseado em Algoritmos Genéticos e Análise de Recência, Freqüência e Valor. Rio de Janeiro: UFRJ/IM/NCE, 2005.

Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro / Instituto de Matemática / Núcleo de Computação Eletrônica, Rio de Janeiro, 2005. Orientador: Antonio Juarez Alencar, D.Phil.

- 1 - Análise de Recência, Freqüência e Valor
- 2 - Algoritmos Genéticos
- 3 - Otimização
- 4 - Marketing Direto
- 5 - Engenharia de Marketing

À **Ligia**, minha esposa que sempre foi meu grande apoio nos momentos mais difíceis. Aos meus pais **Sonia e Nilton** e meu irmão **Fernando**, uma família que não poderia ser melhor .

Agradecimentos

Agradeço à Ligia que sempre me apoiou e me incentivou a continuar caminhando em busca dos meus objetivos apesar de todas as dificuldades.

Ao meu orientador, Professor Juarez, que em todas as conversas e interações sempre permitiu o meu crescimento pessoal e intelectual.

Aos amigos e colegas William, Cláudio e Lívia, pelos incentivos e questionamentos que possibilitaram o enriquecimento contínuo desta tese.

Resumo

RIBEIRO, Eduardo Martins. **Um Modelo Para Otimização de Campanhas de Marketing Baseado em Algoritmos Genéticos e Análise de Recência, Frequência e Valor.** Orientador: Antonio Juarez Alencar. Rio de Janeiro: UFRJ/IM/NCE, 2005. Dissertação (Mestrado em Informática) .

A análise de recência, frequência e valor vem sendo utilizada com sucesso para prever o comportamento dos consumidores em diferentes tipos de campanha de marketing ao redor do mundo. Na análise RFV, consumidores que compraram bens e serviços em um determinado período são agrupados em segmentos de acordo com a data de sua última compra, número de compras no período e seu valor total. Esta tese apresenta como o número de segmentos “RFV” e período de tempo ótimos podem ser determinados. Variações diferentes de análise RFV são consideradas no processo de otimização.

Abstract

RIBEIRO, Eduardo Martins. **Um Modelo Para Otimização de Campanhas de Marketing Baseado em Algoritmos Genéticos e Análise de Recência, Frequência e Valor.** Thesis Supervisor: Antonio Juarez Alencar. Rio de Janeiro: UFRJ/IM/NCE, 2005. Dissertação (Mestrado em Informática) .

The recency, frequency and monetary analysis has been successfully used to predict customer behavior in many different types of marketing campaign all over the world. In RFM analysis customer who have purchased goods and services over a certain period of time are grouped into segments, according to the date of their last purchase, number of purchases within the period and their total value. This thesis shows how the optimum period of time and number of “RFM” segments can be reached. Different variants of RFM analysis are considered in the optimization process.

LISTA DE ILUSTRAÇÕES

Gráfico I: Crescimento médio anual estimado das vendas geradas por marketing direto para o período 2001-2006.

Gráfico II: Crescimento das vendas geradas por marketing direto no Brasil.

Figura I: Passos para o Teste de Mercado.

Figura II: Passos para o teste de mercado com Segmentação.

Figura III: Segmentos de recência da base de clientes.

Figura IV: Segmentos de freqüência da base de clientes.

Figura V: Segmentos de valor monetário da base de clientes.

Figura VI: Ordenação e divisão das células RFV.

Tabela I: Dados fictícios de clientes para a classificação RFV.

Tabela II: Dados de clientes ordenados de acordo com a recência.

Tabela III: Dados de clientes classificados de acordo com a recência.

Tabela IV: Dados de clientes classificados de acordo com a recência e freqüência.

Tabela V: Clientes classificados de acordo com a recência, freqüência e valor.

Figura VII: População de ratos.

Figura VIII: População de ratos em ambiente com novos predadores.

Figura IX: Operadores utilizados no processo.

Figura X: Processo de transformação da população.

Figura XI: *Crossover* entre organismos.

Figura XII: *Crossover* entre organismos. Um dos *offsprings* sofre mutação.

Figura XIII: Representação gráfica da função $g(x)$.

Tabela VI: População inicial e resultado das funções.

Tabela VII: Organismos da população em t_0 .

Tabela VIII: Organismos da população em t_1 .

Tabela IX: Organismos da população em t_{n+1} .

Tabela X: Organismos da população em t_m .

Tabela XI: Destaques Financeiros da Empresa X.

Tabela XII: *Layout* da base de transações.

Tabela XIII: *Layout* da base de resultado da campanha.

Figura XIV: Pasta “Parâmetros” na planilha do RFV Otimizado.

Figura XV: Pasta “Registros no Período” na planilha do RFV Otimizado.

Figura XVI: Pasta “Campanha” na planilha do RFV Otimizado.

Figura XVII: Pasta “Campanha” na planilha do RFV Otimizado.

Figura XVIII: Pasta “Resultado” na planilha do RFV Otimizado.

Figura XIX: Controle das operações do Evolver.

Figura XX: Módulo de configurações do Evolver.

Figura XXI: Opções das células ajustáveis.

Figura XXII: Módulo de opções do Evolver.

Figura XXIII: Relatórios do Evolver.

Figura XXIV: Configuração do modelo no Evolver.

Tabela XIV: Descrição dos parâmetros utilizados nas células ajustáveis.

Tabela XV: Descrição das opções utilizadas no modelo.

Figura XXV: Opções do modelo no Evolver.

Tabela XVI: Estatísticas da campanha de marketing direto da Empresa X.

Tabela XVII: Parâmetros do RFV.

Tabela XVIII: Parâmetros do RFV Otimizado.

Tabela XIX: Comparativos dos resultados da campanha de marketing direto.

Figura XVI: Comparativo de lucro da campanha de marketing direto.

SUMÁRIO

CAPÍTULO 1. INTRODUÇÃO	3
CAPÍTULO 2. MARKETING DIRETO E TESTE DE MERCADO	7
2.1. MARKETING DIRETO	7
2.2. TESTE DE MERCADO	10
2.2.1. DEFINIÇÃO DOS OBJETIVOS DA CAMPANHA	11
2.2.2. SELEÇÃO DO PÚBLICO-ALVO (POPULAÇÃO).....	11
2.2.3. EXTRAÇÃO DE AMOSTRA PARA TESTE.....	12
2.2.4. REALIZAÇÃO DA OFERTA	14
2.2.5. MEDIÇÃO E ANÁLISE DO RESULTADO.....	15
2.2.6. AJUSTE DA OFERTA.....	16
2.2.7. <i>ROLL-OUT</i> DA CAMPANHA	16
2.3. SEGMENTAÇÃO DE CLIENTES	17
2.3.1. SEGMENTAÇÃO DO PÚBLICO-ALVO	18
2.3.2. MEDIÇÃO E ANÁLISE DE RESULTADOS	19
2.3.3. SELEÇÃO DE SEGMENTOS DO PÚBLICO-ALVO	20
2.4. ANÁLISE RFV	20
2.4.1. CODIFICAÇÃO DA BASE DE CLIENTES NA DIMENSÃO DE RECÊNCIA	22
2.4.2. DIMENSÕES DE FREQUÊNCIA E VALOR MONETÁRIO	23
2.4.3. COMBINAÇÃO DOS ÍNDICES R, F E V.....	24
2.4.4. APLICAÇÃO DA ANÁLISE RFV.....	25
2.4.5. EXEMPLO PRÁTICO.....	25
2.5. CONFIGURAÇÃO DA ANÁLISE RFV.....	27
CAPÍTULO 3. ALGORITMOS GENÉTICOS	31
3.1. INTRODUÇÃO	31
3.2. BREVE HISTÓRICO	31
3.3. COMO OS ALGORITMOS GENÉTICOS FUNCIONAM	33
3.3.1. UM EXEMPLO DARWINIANO.....	34
3.3.2. ALGORITMOS GENÉTICOS COMPUTACIONAIS	35
3.3.3. DETALHANDO O OPERADOR DE SELEÇÃO.....	38
3.3.4. <i>DETALHANDO O OPERADOR DE CROSSOVER</i>	39
3.3.5. DETALHANDO O OPERADOR DE MUTAÇÃO.....	41
3.3.6. DETALHANDO O OPERADOR DE INVERSÃO	42
3.3.7. DETALHANDO A FUNÇÃO DE APTIDÃO	42
3.3.8. EXEMPLO COMPUTACIONAL.....	43
3.4. POSSÍVEIS APLICAÇÕES DOS ALGORITMOS GENÉTICOS.....	50
3.4.1. OTIMIZAÇÃO.....	50
3.4.2. GERAÇÃO AUTOMÁTICA DE PROGRAMAS.....	51
3.4.3. INTELIGÊNCIA COMPUTACIONAL.....	51
3.4.4. ECONOMIA.....	52
3.4.5. MARKETING	52
3.4.6. OUTRAS ÁREAS DE CONHECIMENTO	52
3.5. APLICAÇÃO DE ALGORITMOS GENÉTICOS À ANÁLISE RFV NA CONSTRUÇÃO DE UM MODELO DE OTIMIZAÇÃO.....	52

CAPÍTULO 4. FERRAMENTAS E MÉTODO DE PESQUISA.....	54
4.1. INTRODUÇÃO	54
4.2. A EMPRESA X.....	54
4.3. CAMPANHA UTILIZADA	56
4.4. BASE DE DADOS	57
4.5. IMPLEMENTAÇÃO DO RFV OTIMIZADO	58
4.5.1. UTILIZANDO O MS EXCEL	58
4.5.1.1. PASTA DE PARÂMETROS	59
4.5.1.2. PASTA DE REGISTROS DE TRANSAÇÃO	60
4.5.1.3. PASTA DE REGISTROS NO PERÍODO.....	61
4.5.1.4. PASTA DE CAMPANHA	62
4.5.1.5. PASTA DE RESULTADO	64
4.5.2. UTILIZANDO O EVOLVER FOR EXCEL.....	66
4.5.2.1. PORQUE UTILIZAR EVOLVER.....	67
4.5.2.2. COMO O EVOLVER FUNCIONA	70
4.5.3. CRIANDO UM MODELO DE OTIMIZAÇÃO RFV NO EVOLVER.....	73
4.5.3.1. CONFIGURANDO AS CÉLULAS AJUSTÁVEIS.....	74
4.5.3.2. CONFIGURANDO AS OPÇÕES DO EVOLVER	75
CAPÍTULO 5. ANÁLISE DOS RESULTADOS.....	78
5.1. INTRODUÇÃO	78
5.2. RESULTADO REAL DA CAMPANHA	78
5.3. RESULTADO DA CAMPANHA COM ANÁLISE RFV	79
5.4. RESULTADO DA CAMPANHA COM RFV OTIMIZADO.....	80
5.5. COMPARAÇÕES	81
CAPÍTULO 6. CONCLUSÕES	84
6.1. DISCUSSÃO	84
6.2. CONCLUSÕES	89
6.3. SUGESTÕES DE NOVAS PESQUISAS	90
REFERÊNCIAS BIBLIOGRÁFICAS	92

Capítulo 1. INTRODUÇÃO

Nas últimas duas décadas o marketing direto vem se tornando um dos recursos de marketing mais utilizados no Brasil e no mundo. As vendas geradas a partir de seus mecanismos vêm crescendo rapidamente (KOTLER, 2000, p.668), só em 2001 o resultado das vendas de marketing direto nos Estados Unidos superou 1,8 trilhões de dólares. Enquanto a previsão do crescimento geral das vendas de produtos e serviços nos Estados Unidos é de 4,8% para o período 2001-2006, a previsão do crescimento das vendas geradas por marketing direto no mesmo período é de 8,3% (WIENTZEN, 2002).

Segundo a ABEMD, Associação Brasileira de Marketing Direto, mais de 15 bilhões de dólares por ano são gastos nos Estados Unidos apenas com mala direta (ABEMD, 2001). A previsão para gastos com e-mail marketing em 2003 nos Estados Unidos ultrapassava 4,5 bilhões de dólares (EMARKETER, 2000). Somando-se todos os recursos do marketing direto (mala-direta, catálogo, e-mail marketing, tele-marketing, etc.), estes investimentos chegaram a US\$ 196 bilhões em 2001 com estimativa de atingir US\$ 269 bilhões em 2006 (WIENTZEN, 2002).

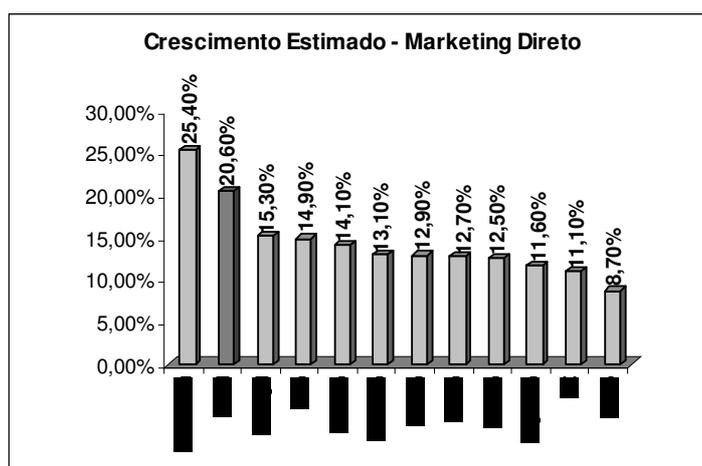


Gráfico I: Crescimento médio anual estimado das vendas geradas por marketing direto para o período 2001-2006 (WIENTZEN, 2002)

Fora dos Estados Unidos, em países que já utilizam o marketing direto, a estimativa para os próximos anos também é de crescimento. No Gráfico I apresentamos as estimativas fornecidas por Wientzen (2002) para o crescimento médio anual de vendas geradas a partir de campanhas de marketing direto para o período de 2001-2006.

No Brasil, ainda segundo Wientzen (2002), as vendas geradas por marketing direto vêm crescendo de forma contínua nestes últimos anos. O crescimento pode ser acompanhado no Gráfico II.

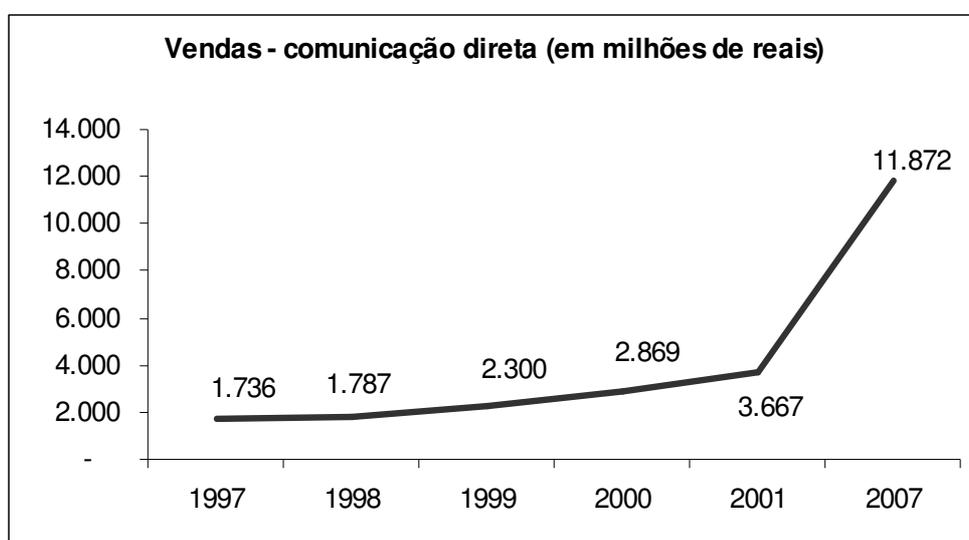


Gráfico II: Crescimento das vendas geradas por marketing direto no Brasil.

Apesar do volume de recursos e esforços investidos, muitas campanhas de marketing direto falham com resultados desastrosos para as organizações e profissionais envolvidos. Hughes (p.335, 2000) chama a atenção para este fato através de vários exemplos de campanhas de marketing direto de alto investimento que falharam. Entre elas podem ser citadas o “*Quaker Direct Project*”, concebido pela *Quaker Oaks Company*, com prejuízos já após o primeiro envio de comunicações dirigidas, o “*Barbie Pink Stamps Club*”, criado pela *Mattel Corporation*, que apesar de ser um projeto de longo prazo foi cancelado nos primeiros três anos e o “*Citicorp’s Reward America*”, realizado pelo gigante financeiro *Citicorp*, cancelado com prejuízos na ordem de 180 milhões de dólares.

Alan Rosenspan (1997) defende em seu artigo que em todas as campanhas de marketing direto há um certo nível de falha. Mesmo as campanhas premiadas, com retorno acima do comum, falham ao obter uma resposta positiva de apenas um pequeno subconjunto do público-alvo.

Quanto maior for o subconjunto que responder positivamente, efetuando a compra de algum produto ou serviço, maior o resultado financeiro gerado pela campanha de marketing direto. Por isso, é natural que os profissionais envolvidos nas campanhas de marketing direto façam esforços contínuos de aumento do subconjunto de pessoas que respondem positivamente às ofertas que lhes são feitas.

Para tanto, os profissionais de marketing direto procuram ajustar as ofertas que suas organizações fazem, o mais fielmente possível, à demanda do mercado. Este ajuste procura obter um resultado mais positivo das campanhas de marketing direto oferecendo maior satisfação ao consumidor, atingida através da diferenciação da oferta feita pela organização em relação aos concorrentes, a mesma satisfação por um preço menor ou ambos (ROCHA e CHRISTENSEN, 1999).

O ajuste da oferta à demanda pode ser feito de duas formas distintas. Uma delas é através de um processo de tentativa e erro baseado em intuição e experiências progressas, a outra se dá através de testes de mercado, onde os profissionais de marketing procuram inferir estatisticamente os produtos e serviços que o mercado deseja comprar, assim como as condições adequadas de venda (AAKER, KUMAR e DAY, 2001).

Neste trabalho apresentamos um modelo para otimização do resultado das campanhas de marketing direto baseado em uma das técnicas mais populares de teste de mercado neste segmento, isto é, a análise de recência, frequência e valor, Análise RFV (Hughes, 2000). O modelo de otimização apresentado é baseado em técnicas de inteligência computacional fazendo amplo uso de algoritmos genéticos (MICHELL, 1998).

Ao longo deste trabalho estarão sendo respondidas as seguintes perguntas:

- Como o modelo que apresentamos maximiza os resultados de uma campanha de marketing direto?
- Por que esforços devem ser empreendidos para otimizar a Análise RFV?
- Por que a Análise RFV deve ser otimizada?
- Como esta otimização pode ser realizada?
- Como a Análise RFV Otimizada pode ser utilizada para tornar uma campanha de marketing direto mais lucrativa?
- Como a Análise RFV Otimizada pode ser utilizada para determinar, com maior precisão, o perfil dos clientes que deverão fazer parte da campanha?
- Qual o impacto para as estratégias corporativas e de marketing de um modelo que maximiza os resultados das campanhas de marketing direto?
- Como um modelo de otimização de resultado de campanha de marketing direto beneficia os clientes de uma organização e a sociedade em geral?

Capítulo 2. MARKETING DIRETO E TESTE DE MERCADO

2.1. MARKETING DIRETO

O marketing direto, como atividade, remonta há mais de dois mil anos antes de Cristo, quando comerciantes do Golfo Pérsico enviavam suas mercadorias em barcos para serem vendidos diretamente aos seus clientes. Já naquela época os comerciantes tinham conhecimento de que pequenos ajustes no formato, conteúdo e preço das mercadorias favoreciam substancialmente as vendas (RAPHAEL, 2000).

Infelizmente, muitas das práticas empíricas de marketing direto adquiridas à duras penas através dos anos acabaram ficando esquecidas durante a Revolução Industrial (século XVIII). Neste período, a demanda por produtos se manteve maior que a oferta e, portanto, o cliente não possuía grandes poderes de negociação (GUIMARÃES e BRISOLA, 2003, p.2). Em outras palavras, tudo o que era fabricado era vendido.

Foi somente no início do século XIX que as práticas do marketing direto voltaram a ficar em evidência com a venda porta-a-porta. Um dos pioneiros destas práticas foi *Eli Terry*, que vendia seus relógios customizados no estado de *Connecticut*, Estados Unidos.

A partir de *Eli Terry* as praticas de marketing direto se intensificaram no comércio varejista da época. No final do século XIX, por exemplo, alguns comerciantes ficaram famosos anunciando suas mercadorias com o uso de tambores enquanto viajavam por várias cidades, ficando conhecidos como *drummers*¹ (RAPHAEL, 2000).

¹ Percussionistas

A prática de se anunciar produtos com auxílio de tambores acabou por dar origem aos catálogos de produtos para venda direta, que eram colocados nas caixas de correio de clientes potenciais. Os irmãos *Butler* foram pioneiros no uso desta técnica de venda, lançando em 1877 um dos primeiros catálogos, chamado *Our Drummer*².

Paradoxalmente, o sucesso das vendas por catálogo acabou se tornando uma grande ameaça para os vendedores viajantes, que tinham inspirado o próprio surgimento destes catálogos. Tanto que muitos contratavam jovens nas cidades em que visitavam para recolher das caixas de correio os catálogos da *Montgomery Ward*, um grande usuário das práticas de marketing direto nos anos 1900s.

Ao longo do século XX, o marketing direto foi evoluindo, adquirindo novos adeptos e experimentando novas ferramentas e canais. Não é mais usual que o vendedor faça contato porta-a-porta como fazia *Eli Terry*. Neste século, a oferta de marketing direto chega aos clientes e *prospects*³ através dos correios, telefone e correio eletrônico (RAPHAEL, 2000).

Conforme profetizado pelo guru do marketing direto *Lester Wunderman* há um quarto de século, o consumidor não precisa mais sair de casa nos dias de hoje em busca das mercadorias que deseja. Ao invés disso, ele pode permanecer em casa, de onde é possível comprar quase tudo aquilo que deseja. *Wunderman* definiu este novo paradigma do marketing direto do século XX como uma experiência de “implosão” contrastando com a experiência do tipo “explosão”, praticada no passado (SHETH e PARVATIVAR, 1995).

A popularização do marketing direto foi impulsionada pelo desenvolvimento tecnológico e pela crescente demanda por produtos e serviços com graus de personalização cada vez maiores. Raphael (2000) e Sheth e Parvatiyar (1995) fazem uma análise detalhada da história do marketing direto.

² Nosso percursorista.

³ Clientes potenciais.

Apesar da longa história do marketing direto, ainda hoje existe alguma controvérsia relacionada à definição desta atividade. Segundo a Associação Brasileira de Marketing Direto, “por Marketing Direto entende-se a disciplina de Marketing cuja comunicação se utiliza de uma ou mais mídias para obter uma resposta ou transação mensurável junto a públicos específicos, ou simplesmente gerar uma ação de relacionamento que gere encantamento junto ao público visado” (ABEMD, 2003).

Outra associação, a Associação Brasileira de Marketing & Negócios, também atuando na indústria do marketing, define Marketing Direto como “conjunto de atividades de promoção de vendas e de comercialização em que se desenvolve um relacionamento direto entre a empresa (que oferece um produto ou serviço) e o consumidor” (ABMN, 2002).

Nos Estados Unidos, a *Direct Marketing Association*⁴ define o Marketing Direto como “um sistema interativo de marketing que usa uma ou mais mídias de propaganda para obter uma resposta mensurável e / ou uma transação em qualquer lugar” (DMA *apud* KOTLER, 2000).

Já Hughes, em seu livro *Strategic Database Marketing* (2000, p.432), diz ser o Marketing Direto “qualquer sistema de marketing em que espera-se que o consumidor responda diretamente para o anunciante (ao invés de ir a uma loja). Marketing Direto pode utilizar mídias de comunicação de massa, correio, telefone ou a Internet”.

No escopo desta tese entendemos que o marketing direto é uma disciplina do marketing que compreende todas as atividades de interação com o cliente ou *prospect* através de um canal direto de comunicação. A interação pode ter o objetivo de gerar alguma transação mensurável ou de apoiar o relacionamento com o cliente ou *prospect*. Como canal direto de comunicação consideramos qualquer meio em que um cliente ou *prospect*, identificado unicamente, recebe uma mensagem da organização praticante do marketing direto. Esta definição é, na verdade, uma consolidação das definições acima mencionadas.

⁴ Associação Americana de Marketing Direto.

2.2. TESTE DE MERCADO

Uma das formas mais utilizadas pelos profissionais de marketing direto para ajustar a oferta à demanda dos clientes e *prospects* é o teste de mercado. No teste de mercado busca-se inferir estatisticamente os produtos e serviços que o mercado deseja comprar e as condições de compra mais adequadas (AAKER, KUMAR e DAY, 2001).

Hughes (2000) propõe a utilização do teste de mercado em conjunto com a análise de recência, freqüência e valor e Sullivan (2000) defende a utilização das técnicas de teste de mercado para construir campanhas de alta receita com retorno de investimento. Em um teste de mercado, algumas etapas precisam ser cumpridas na busca de se obter uma estimativa confiável do resultado de uma campanha de marketing. Estes passos estão representados na Figura I e são detalhados a seguir.

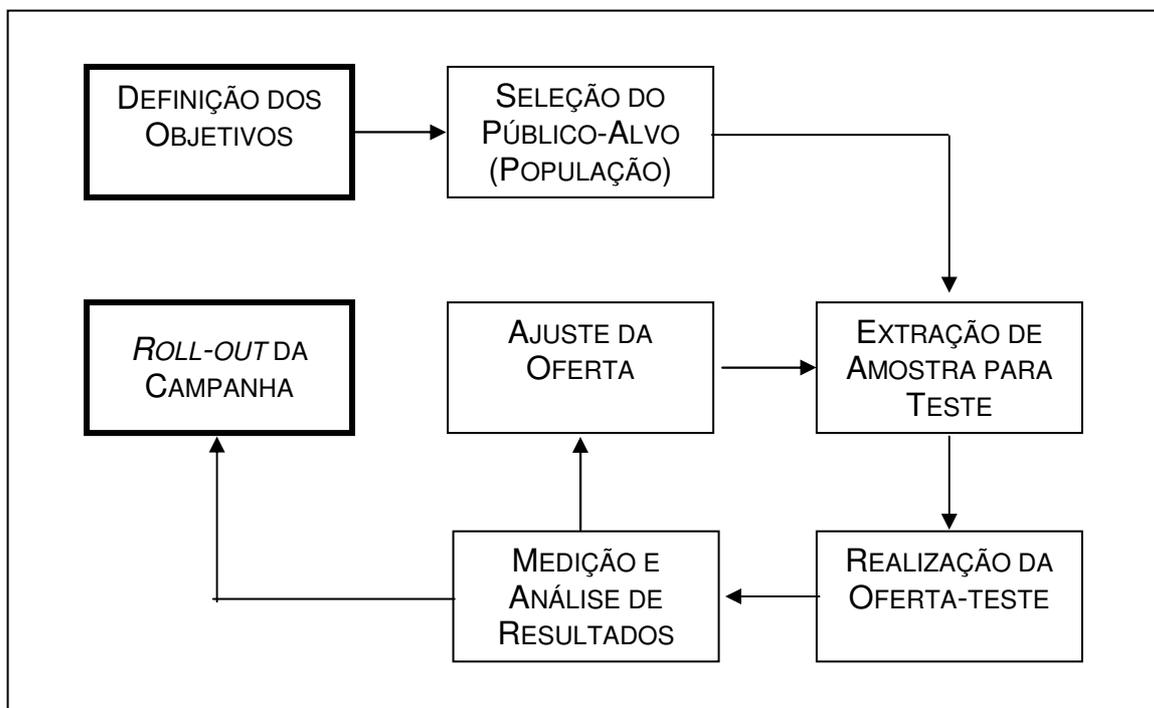


Figura I: Passos para o Teste de Mercado

2.2.1. DEFINIÇÃO DOS OBJETIVOS DA CAMPANHA

Para que o teste de mercado seja efetivo, a organização praticante de marketing direto deve ter os objetivos da campanha claramente definidos. Estes objetivos devem, ainda, ser mensuráveis para que possam ser utilizados na fase de “Medição e Análise de Resultados”. Será com base nestas medidas que os profissionais envolvidos analisarão se o resultado estimado da campanha atingirá os objetivos e, portanto, deve ser executada.

Por exemplo, se uma organização deseja simplesmente que uma campanha de marketing direto propicie um certo retorno financeiro, então ela deve especificar que retorno financeiro a campanha deverá propiciar. Por outro lado, se a organização deseja incrementar as vendas de um determinado produto (ou linha de produtos) em um certo percentual, este percentual deve ser estabelecido claramente. Outros objetivos possíveis são o aumento da quantidade de solicitações de informações sobre um determinado produto ou serviço, aumentar a circulação de pessoas nas lojas físicas, o número de visitas no *site* da organização, reativar um certo volume de clientes etc.

Um dos objetivos mais comuns de uma campanha de marketing é o de atingir ou ultrapassar o ponto de *breakeven*. Hughes (2000, p.431) define *breakeven* como o ponto em que “o lucro de vendas depois dos impostos é exatamente igual aos custos de comunicação com os clientes”. Em outras palavras, é o volume de vendas que cobre exatamente os custos de comunicação da campanha de marketing direto. Um volume inferior determina perdas e um volume superior ganhos financeiros.

2.2.2. SELEÇÃO DO PÚBLICO-ALVO (POPULAÇÃO)

Segundo Bitran e Mondschein (1997), uma das decisões mais importantes que precisa ser tomada nas atividades de marketing direto é quem vai receber as comunicações. Ou seja, que clientes e *prospects* serão alvos de uma campanha de marketing direto.

Atualmente muitas organizações possuem políticas de comunicação que utilizam regras rígidas de *opt-in*⁵ para selecionar quais clientes devem fazer parte do público-alvo de uma campanha de marketing direto (EMARKETER, 2000). Por exemplo, apenas os clientes que explicitamente autorizaram o recebimento de comunicações nos últimos 12 meses podem ser selecionados.

Em adição ao *opt-in* é natural que os responsáveis pelas campanhas de marketing direto realizem também uma pré-seleção baseada em intuição e experiências pregressas. Por exemplo, se a oferta da campanha de marketing direto for de produtos ou serviços para crianças, o público-alvo selecionado poderá ser de clientes com filhos ou netos na faixa-etária adequada à oferta. Por outro lado, se os produtos ou serviços ofertados irão estar disponíveis apenas em uma certa região geográfica, por exemplo, no estado do Rio de Janeiro, é razoável que apenas os cliente que residam na região geográfica em questão sejam selecionados para participarem da campanha.

Consideradas as políticas e pré-seleções, um grupo de clientes é separado como candidatos ao recebimento da comunicação. Este grupo será a população da qual será extraída a amostra para a comunicação da oferta-teste.

2.2.3. EXTRAÇÃO DE AMOSTRA PARA TESTE

Uma amostra estatisticamente aleatória deve ser retirada da população. Se a amostra for retirada corretamente, então ela irá representar, tão fielmente quanto possível, os perfis demográficos, psicográficos e comportamentais da população original. Oferecendo, portanto, aos profissionais da área uma excelente oportunidade para medir e estimar o comportamento da população com um custo de comunicação reduzido. Tamanhos típicos de amostra para campanhas de marketing direto variam entre 200 e 3000 indivíduos.

Triola (1998, p.144) chama a atenção para a importância de se extrair corretamente a amostra já que “dados coletados de forma imprecisa ou

⁵ Inclusão voluntária.

descuidada podem ser totalmente destituídos de valor, mesmo que a amostra seja suficientemente grande”.

Para que os dados da amostra sejam coletados de forma correta cada elemento da população deve ter a mesma chance de ser selecionado. Além disso a amostra deve possuir um tamanho mínimo para permitir que as incertezas associadas às estimativas que ela propiciar sejam mantidas dentro de um patamar aceitável para a organização interessada na execução da campanha.

Amostras muito pequenas tendem a propiciar estimativas com alto grau de incerteza a um custo extremamente reduzido, enquanto amostras maiores tendem a propiciar estimativas mais precisas com custo maior.

Pode-se utilizar a pesquisa eleitoral para escolha do Presidente da República no Brasil como exemplo de extração de amostra para estimativa de resultado. Um grupo de cerca de pouco mais de 2000 eleitores é, em geral, entrevistado pelos institutos de pesquisa para avaliar a intenção de voto. A partir da análise dos resultados propiciados pela amostra gera-se uma estimativa do resultado geral da eleição com uma margem de erro de aproximadamente dois pontos percentuais, levando-se em consideração um intervalo de confiança de 95%.

Um exemplo de amostra muito pequena seria perguntar a intenção de voto a apenas trinta pessoas e estimar o resultado da eleição a partir desta amostra. Neste caso, o erro percentual associado às estimativas do resultado da eleição seria de aproximadamente 15%. Em decorrência, candidatos com 30% e 44% das intenções de voto estariam tecnicamente empatados. Vale mencionar que o instituto de pesquisa IBOPE (2004) declara ter entrevistado mais de 2.000 pessoas nas pesquisas que antecederam a eleição para Presidente da República em 2002.

Um exemplo de amostra coletada de forma imprecisa seria perguntar a intenção de votos a 2.000 pessoas que compareceram à convenção do partido

político PSDB. Muito provavelmente o resultado das análises da pesquisa revelaria, de forma incorreta, que o candidato do partido ganharia as eleições por uma larga margem de votos. Para evitar este tipo de problema, o IBOPE (*Op.cit*) entrevistou eleitores em 143 municípios brasileiros. Triola (1998) apresenta um estudo detalhado sobre extração e cálculo do tamanho mínimo de amostras.

2.2.4. REALIZAÇÃO DA OFERTA

Nesta fase, os clientes que fazem parte da amostra recebem da organização uma oferta através de uma comunicação dirigida⁶. Esta comunicação dirigida, conforme citado anteriormente, pode utilizar os recursos de mala-direta, catálogo, e-mail marketing, tele-marketing etc.

É importante observar que somente se as mesmas condições estabelecidas na execução do teste forem mantidas no *roll-out*⁷ da campanha, é que se poderá garantir que os resultados obtidos do teste espelharão com precisão os resultados da campanha. Neste sentido, se o gerente de marketing deseja utilizar os resultados do teste para estimar os resultados da campanha, então ele não deve alterar o *mix* de marketing, ou seja a oferta, a linguagem e o formato da comunicação, as condições de pagamento etc., no *roll-out* da campanha.

Infelizmente, no dia-a-dia do marketing direto o *roll-out* de uma campanha nem sempre tem o mesmo resultado dos testes. Isso ocorre, principalmente, porque as condições de mercado se modificaram substancialmente entre a realização da oferta-teste e o *roll-out* da campanha ou os melhores recursos foram aplicados durante o teste para justificar a realização da campanha, não sobrando bons recursos (oferta, público-alvo etc.) para a campanha propriamente dita (HUGHES, 2000, p.132).

⁶ Comunicação dirigida e comunicação direta são sinônimos encontrados na literatura com o mesmo significado, que é a comunicação enviada para um cliente ou *prospect* em uma campanha de marketing direto. Nesta tese, optamos por utilizar o termo comunicação dirigida.

⁷ Execução completa de uma campanha de marketing.

2.2.5. MEDIÇÃO E ANÁLISE DO RESULTADO

A forma de medir o resultado do teste, assim como da campanha, pode variar dependendo do objetivo que se deseja atingir, a saber: incremento de venda ou intensificação de relacionamento. O fundamental é que o seu retorno seja mensurável através de alguma transação.

Se a campanha tiver o objetivo de aumentar as vendas, então as transações de venda motivadas pela campanha serão o melhor indicador de sucesso e, portanto, precisam ser identificadas. Por outro lado, se a campanha tiver o objetivo de intensificar o relacionamento com os clientes então outros indicadores, tais como visita à loja, solicitação de informações no *call-center*, degustação dos produtos ou serviços a serem adquiridos etc, serão mais apropriados para medir o resultado da campanha de marketing direto, devendo também ser identificados.

Portanto a medição do resultado de uma campanha de marketing direto passa pela identificação das transações de venda ou relacionamento que foram motivadas por ela. Em geral, esta identificação é realizada explicitamente solicitando que o cliente informe que campanha motivou a compra de produtos e serviços da organização. Por exemplo, muitas organizações propiciam um desconto promocional à seus clientes somente se eles apresentarem um cupom ou peça publicitária no ato da compra.

Um outro mecanismo utilizado pelas organizações é o da aplicação de uma regra de identificação que considera determinados atributos de uma campanha de marketing direto, tais como: oferta, público-alvo, canal de venda, condições de pagamento, prazo de validade da oferta etc. Por exemplo, se está registrado nos bancos de dados de uma rede de livrarias que um determinado cliente comprou o livro mais recente do Paulo Coelho em uma determinada data e preço, e tudo isso corresponde aos atributos da campanha do Dia dos

Namorados, então a transação de venda é computada nos resultados desta campanha.

2.2.6. AJUSTE DA OFERTA

Após identificadas todas as transações atribuídas a uma campanha, deve-se verificar se os objetivos foram atingidos. Se foram atingidos ou ultrapassados, então o resultado do teste foi positivo e a campanha deve ser executada (*roll-out*). Caso contrário, os profissionais envolvidos podem decidir realizar um novo teste ajustando as condições da oferta ou, em última instância, cancelar a campanha.

Se os profissionais de marketing optarem por realizar o ajuste da oferta, uma nova amostra deve ser extraída e um novo teste realizado. Os passos de extração da amostra, realização da oferta e medição e análise dos resultados se repetem até que o objetivo seja atingido ou a campanha cancelada.

No caso de seguidas atividades de ajuste da oferta, é importante que apenas uma condição (produto, forma de pagamento, mensagem, canal de venda etc.) da campanha seja trocado por execução do teste de mercado, se o praticante deseja ter ciência do motivo da alteração na estimativa do resultado (SULLIVAN, 2000).

2.2.7. ROLL-OUT DA CAMPANHA

No *roll-out* da campanha os clientes e *prospects* que fazem parte do público-alvo e que não participaram das amostras teste recebem a versão final das peças de comunicação com as ofertas da campanha de marketing direto. Neste momento se aplicam algumas restrições apresentadas por Bitran e Mondschein (1997):

- Restrições de Orçamento - os custos das transações resultantes da campanha de marketing direto somados aos custos de produção e

envio das peças de comunicação devem ficar abaixo do orçamento destinado à campanha de marketing direto.

- Restrições de Inventário - as transações geradas a partir da campanha de marketing direto precisam ser iguais ou estar abaixo do nível de inventário de produtos disponíveis para atender a estas transações.
- Restrições das Listas de *Prospects* - limita o número máximo de nomes a serem alugados das listas de *prospects* para o envio de peças de comunicação da campanha de marketing direto ao máximo de nomes disponíveis no mercado para locação.
- Restrições da Base de Clientes - similar à restrição relacionada às listas externas, a quantidade de clientes disponíveis na base interna limitará a quantidade de peças de comunicação que podem ser enviadas na campanha de marketing direto.

Aplicadas as restrições e realizada a campanha de marketing direto, deve-se medir o resultado real da campanha para que seja inferida a qualidade dos testes realizados.

2.3. SEGMENTAÇÃO DE CLIENTES

O teste de mercado, da maneira como foi apresentado no item anterior, trata todo o público-alvo de forma uniforme. Ou seja, todos os clientes e *prospects* fazem parte de um único conjunto e recebem o mesmo tratamento em relação às condições da campanha de marketing direto. Porém, o conjunto de clientes e *prospects* que são atingidos por uma comunicação dirigida é, em geral, formado por indivíduos de diferentes perfis. Cada indivíduo responderá à oferta de uma maneira diferenciada.

Por exemplo, determinados indivíduos não realizarão nenhuma compra embora tenham ficado satisfeitos com o recebimento das peças de comunicação, enquanto outros irão solicitar a remoção de seus nomes do

banco de dados da organização que patrocinou a campanha. Alguns indivíduos realizarão compras de pequeno porte e outros irão realizar compras em um volume maior propiciando um lucro expressivo para a organização em questão.

O objetivo da segmentação de clientes é separar estes indivíduos em grupos homogêneos de acordo com determinadas características. Neste caso, deseja-se dividir os clientes e *prospects* em grupos com um grau semelhante de resposta a uma campanha de marketing direto. Segundo Myers (1996, p.4), o objetivo é “identificar subgrupos de clientes ou *prospects* que, espera-se, responderão de uma maneira similar a um determinado *mix* de marketing”.

Desta forma, é possível medir o retorno do teste separadamente para cada segmento e estimar o retorno destes segmentos no *roll-out* da campanha de marketing direto. Em geral, cada segmento tem índices de resposta diferentes. Parte dos segmentos atingirão o objetivo, enquanto outros terão o retorno estimado abaixo do objetivo estabelecido.

Uma forma, então, de otimizar o resultado da campanha de marketing direto é separar os segmentos que atingirão ou ultrapassarão os objetivos e realizar o *roll-out* apenas com estes segmentos. A implementação deste processo pode ser feita através da inclusão de duas novas etapas no teste de mercado, que são a “Segmentação do Público-Alvo” e a “Seleção Segmentos do Público-alvo”. O processo completo é representado na Figura II.

Estas novas etapas são detalhadas a seguir, as demais são idênticas às apresentadas na Figura I, à exceção da “Medição e Análise de Resultado” e por isso não serão novamente abordadas nesta seção

2.3.1. SEGMENTAÇÃO DO PÚBLICO-ALVO

Depois que um grupo de clientes é separado como candidatos ao recebimento da comunicação, antes que seja extraída a amostra para envio do teste, este grupo passa por um processo de segmentação. A segmentação irá

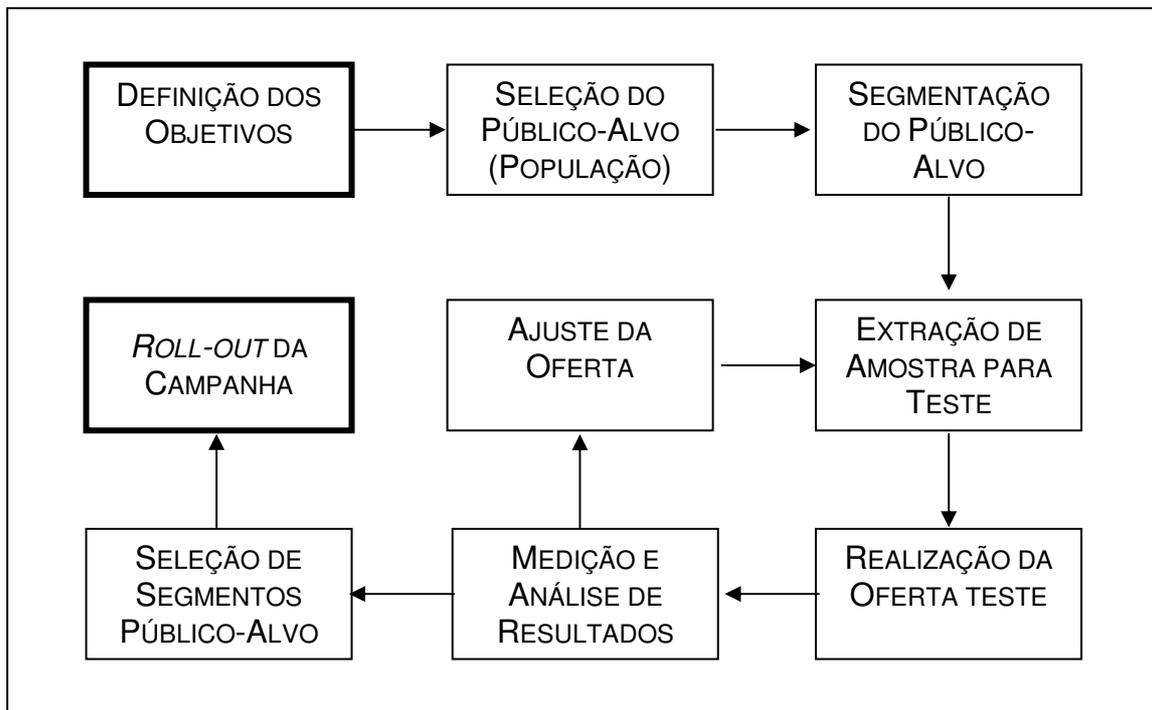


Figura II: Passos para o Teste de Mercado com Segmentação

separar os clientes em grupos homogêneos com base em determinadas características.

Embora existam inúmeras propostas para segmentação de clientes (EVERITT *et al*, 2001; WEDEL *et al*, 1998; MYERS, 1996; SWENSON, 1996), os processos de segmentação mais populares em marketing direto são a segmentação demográfica e a segmentação por comportamento de compra. Isso é corroborado por Hughes (2000, p.105), que afirma existir, basicamente, dois tipos de informações que a organização pode saber de seus clientes: quem eles são (dados demográficos) e o que eles fazem (dados comportamentais). As soluções de segmentação de clientes, em geral, utilizarão estas informações.

2.3.2. MEDIÇÃO E ANÁLISE DE RESULTADOS

Em adição às idéias apresentadas no item “Medição e Análise de Resultados” da Figura I, a medição de resultados com segmentação de clientes deve ser realizada para cada segmento em separado. Cada um destes

resultados será utilizado para estimar o resultado do mesmo segmento na população original caso seja realizada o *roll-out* da campanha.

Nesta etapa, os segmentos que atingirão ou ultrapassarão os objetivos previamente estabelecidos devem ser separados daqueles que não os atingirão. Estes últimos podem sofrer etapas de ajuste da oferta ou simplesmente serem ignorados no *roll-out* da campanha.

2.3.3. SELEÇÃO DE SEGMENTOS DO PÚBLICO-ALVO

Realizadas todas as etapas de ajuste da oferta, para o *roll-out* da campanha apenas os clientes dos segmentos que atingirão ou ultrapassarão os objetivos estabelecidos devem ser selecionados. Desta forma, a organização praticante do marketing direto evita que sejam investidos recursos na realização da campanha para os grupos que, estima-se, não responderão adequadamente à oferta.

2.4. ANÁLISE RFV

Embora existam várias métodos para o agrupamento de clientes em segmentos homogêneos para a previsão e otimização de retorno de campanhas de marketing direto, tais como redes neurais, algoritmos genéticos, árvore de classificação, k-means entre outras (FITZPATRICK, 2001; EVERITT, 2001), o mercado tem optado de forma expressiva pelo uso da análise do comportamento de compra dos clientes nas dimensões de recência, frequência e valor monetário, isto é, Análise RFV (MIGLAUTSCH, 2002; PELTIER, SCHIBROWSKY, SCHULTZ e DAVIS, 2002; RAINEY, 2001; HUGHES, 2000; NICOLAI, 1999; SCHMID, 1998).

Esta análise já vem sendo utilizada com sucesso há muito tempo pelas áreas de marketing de organizações em todo o mundo, tendo sua importância apontada pela primeira vez por *George S. Cullinan* em 1961 (DMA, 2003). De acordo com Hughes (2000, p.105), a Análise RFV já é utilizada há quase 50 anos por profissionais de marketing direto.

Razões para sua popularidade não faltam. Trata-se de um modelo barato e simples, ainda que muito poderoso para prever taxas de respostas das campanhas de marketing direto e, com isso, otimizar o seu resultado (HUGHES, 2000, p.123).

Nicolai (1999, p.194) concorda com Hughes (*Op.cit.*) quando afirma ser o RFV um grande conceito. “Essencialmente, RFV – Recência, Frequência e Valor – é a primeira e ainda a melhor técnica de seleção que qualquer operador de catálogo pode utilizar para encontrar consumidores”. Ainda completa dizendo que “a análise RFV é tão simples quanto poderosa”.

Além disso, a Análise RFV pode ser considerada a base para qualquer modelo de comportamento de clientes. Hughes (2000, p.123) afirma que “qualquer bom modelo de comportamento de clientes sempre inclui recência, frequência e valor monetário”.

Para codificar a base de clientes na análise RFV é necessário possuir três informações básicas sobre as transações realizadas por cada cliente. Estas informações dizem respeito a:

- i) a data da última transação realizada, utilizada para a codificação de recência;
- ii) a quantidade de transações realizadas pelo cliente, utilizada para a codificação de frequência; e
- iii) o valor monetário aplicado pelo cliente no negócio da organização, utilizado para a codificação de valor monetário.

Todas as informações supracitadas são calculadas a partir de um determinado período de tempo, chamado de período de análise, que é arbitrariamente escolhido pelo gerente de marketing.

Durante a pesquisa realizada para desenvolvimento desta tese foram encontradas algumas variações na forma de implementação do RFV (MIGLAUTSCH, 2002; PELTIER, SCHIBROWSKY, SCHULTZ e DAVIS, 2002;

RAINEY, 2001; HUGHES, 2000; NICOLAI, 1999; SCHMID, 1998). Para desenvolvimento do modelo apresentado por este trabalho o método utilizado foi o apresentado por Hughes (2000, p.123).

2.4.1. CODIFICAÇÃO DA BASE DE CLIENTES NA DIMENSÃO DE RECÊNCIA

Para determinar os pontos de corte que dividirão os clientes em diferentes segmentos na dimensão recência, ordena-se a base de clientes a partir da data da última transação. Os clientes com transação mais recente ficam no topo da lista. Estes clientes são, então, divididos em um determinado número de segmentos com o mesmo tamanho.

Tradicionalmente a decisão da quantidade de segmentos é arbitrária e, seguindo as indicações de Hughes (2000, p.106), o mais comum é dividir em cinco partes de tamanhos iguais. Cada segmento recebe um número que indica o grau de recência, sendo que o mais recente recebe o número 5 e o menos recente o número 1, conforme exemplificado na Figura III.

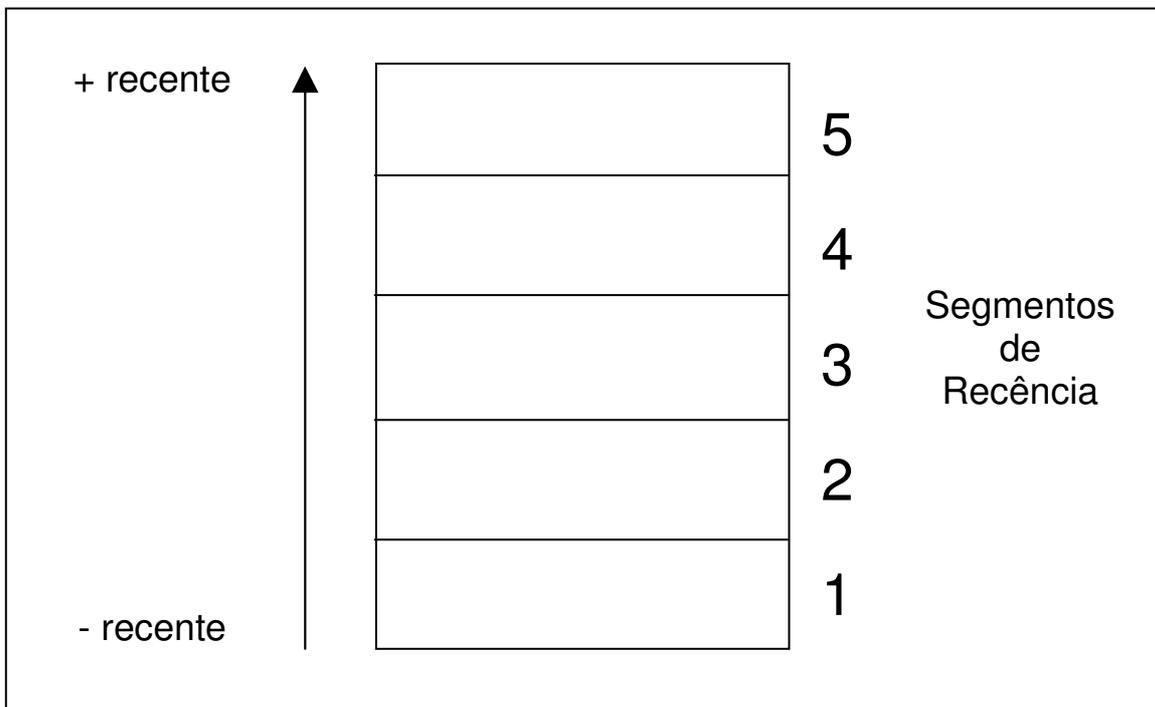


Figura III: Segmentos de recência da base de clientes

2.4.2. DIMENSÕES DE FREQUÊNCIA E VALOR MONETÁRIO

A classificação da base de clientes nas dimensões de frequência e valor monetário segue o mesmo processo apresentado anteriormente para a classificação na dimensão de recência.

Seguindo a estratégia aplicada à recência, a base de dados deve ser ordenada de acordo com a frequência de realização de compras em ordem decrescente (do mais freqüente para o menos freqüente). Os clientes são divididos em um determinado número de segmentos do mesmo tamanho, tipicamente cinco segmentos. Cada segmento recebe um número que indica o grau de freqüência, sendo que o mais freqüente recebe o número 5 e o menos freqüente o número 1. Ver Figura IV.

Em relação ao valor monetário, aplica-se a mesma estratégia da segmentação por freqüência, ordenando-se os clientes de forma decrescente de acordo com o volume de dinheiro destinado a transações com a organização, conforme exemplificado na Figura V.

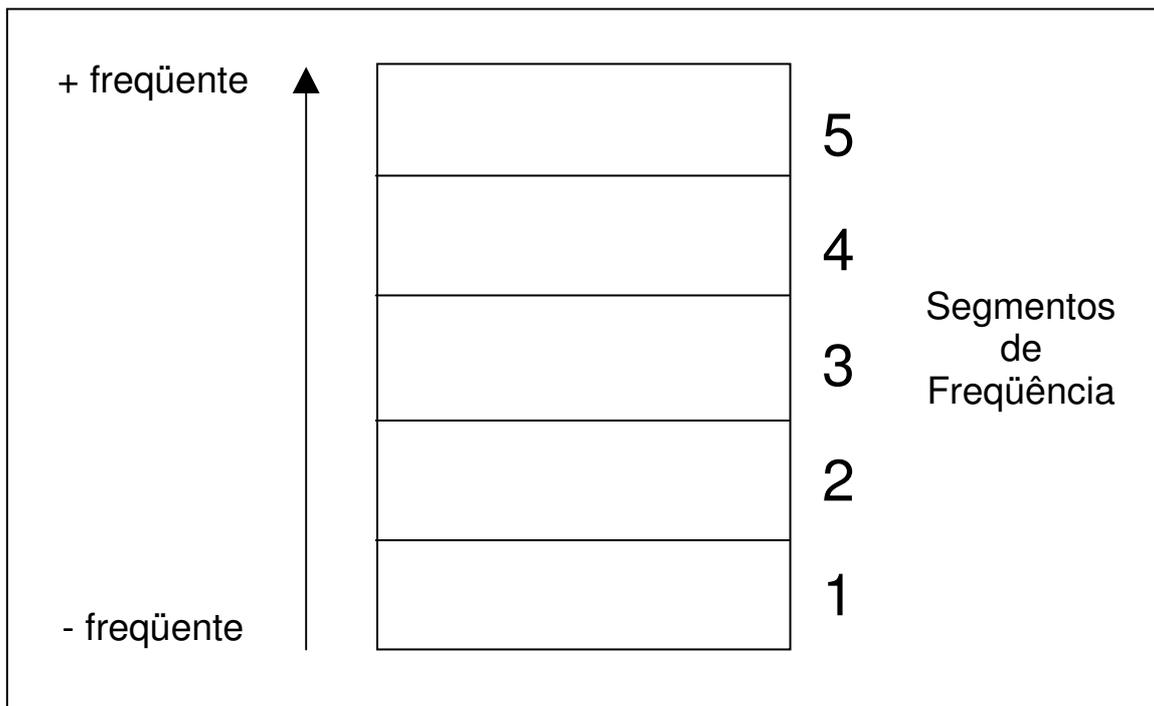


Figura IV: Segmentos de freqüência da base de clientes

2.4.3. COMBINAÇÃO DOS ÍNDICES R, F E V

No final do processo de codificação dos segmentos de recência, frequência e valor monetário, cada cliente da base de dados possui um código RFV. A quantidade de segmentos em cada dimensão irá determinar a quantidade total de segmentos de clientes.

Seguindo o exemplo apresentado por Hughes (*Op.cit*), cada dimensão foi dividida em cinco segmentos. Logo a combinação destes cinco segmentos nas três dimensões para cada cliente gera 125 diferentes grupos ($5 \times 5 \times 5 = 125$). Cada cliente faz parte de um e somente um segmento RFV.

A combinação das classificações nas três dimensões, porém, pode ser feita de maneiras diferentes. Hughes (*Op.cit*) propõe que a ordenação seja realizada primeiro por recência. Depois, cada segmento de recência é ordenado por frequência e devidamente dividido. Segue-se, então, ordenando cada segmento recência-freqüência pelo valor monetário e realiza-se a divisão dos

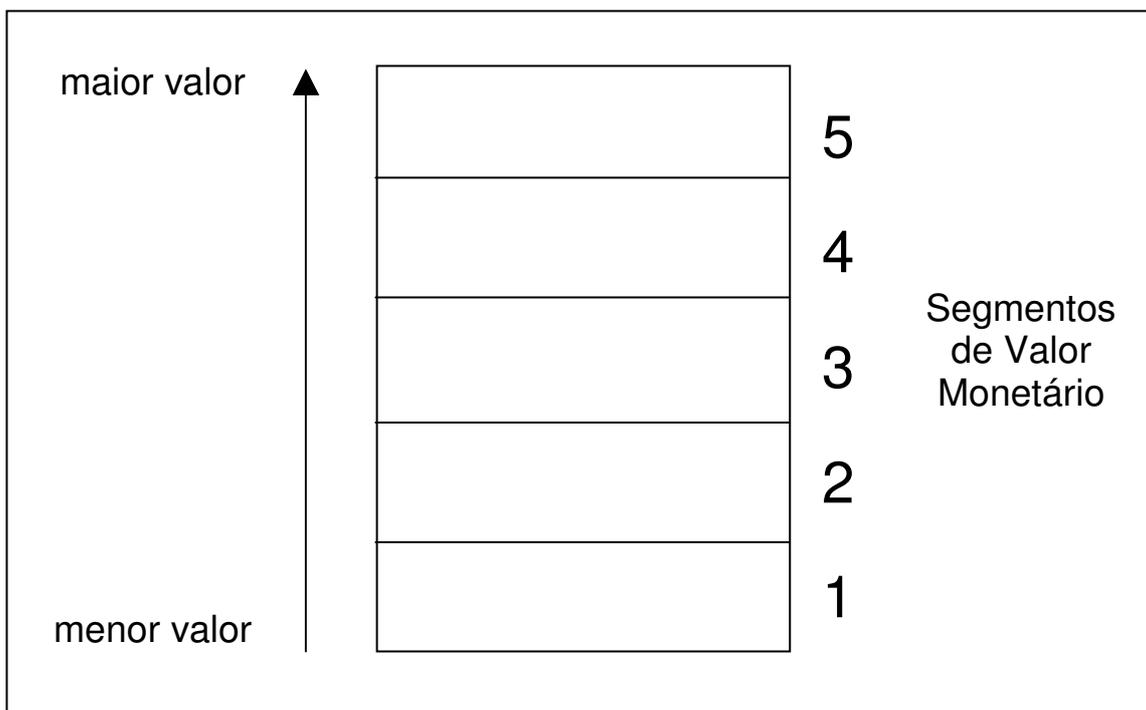


Figura V: Segmentos de valor monetário da base de clientes

segmentos. A Figura VI exemplifica este tipo de ordenação graficamente.

2.4.4. APLICAÇÃO DA ANÁLISE RFV

Conforme já foi citado nesta tese, a Análise RFV possibilita a otimização dos resultados das campanhas de marketing direto. Muitas organizações vêm utilizando esta técnica para gerar ou aumentar os ganhos financeiros das suas campanhas de marketing.

A forma de extrair este benefício da Análise RFV é, em conjunto com o teste de mercado, não enviar comunicações aos grupos que não responderão positivamente à campanha de marketing direto (HUGHES, 2000, p.126). Desta forma o resultado estimado das campanhas de marketing direto serão sempre positivos, já que as campanhas que o teste demonstrar um resultado negativo poderão sofrer ajustes ou serem canceladas.

Schimid e Boyle (1998) sugerem uma quebra adicional na classificação com

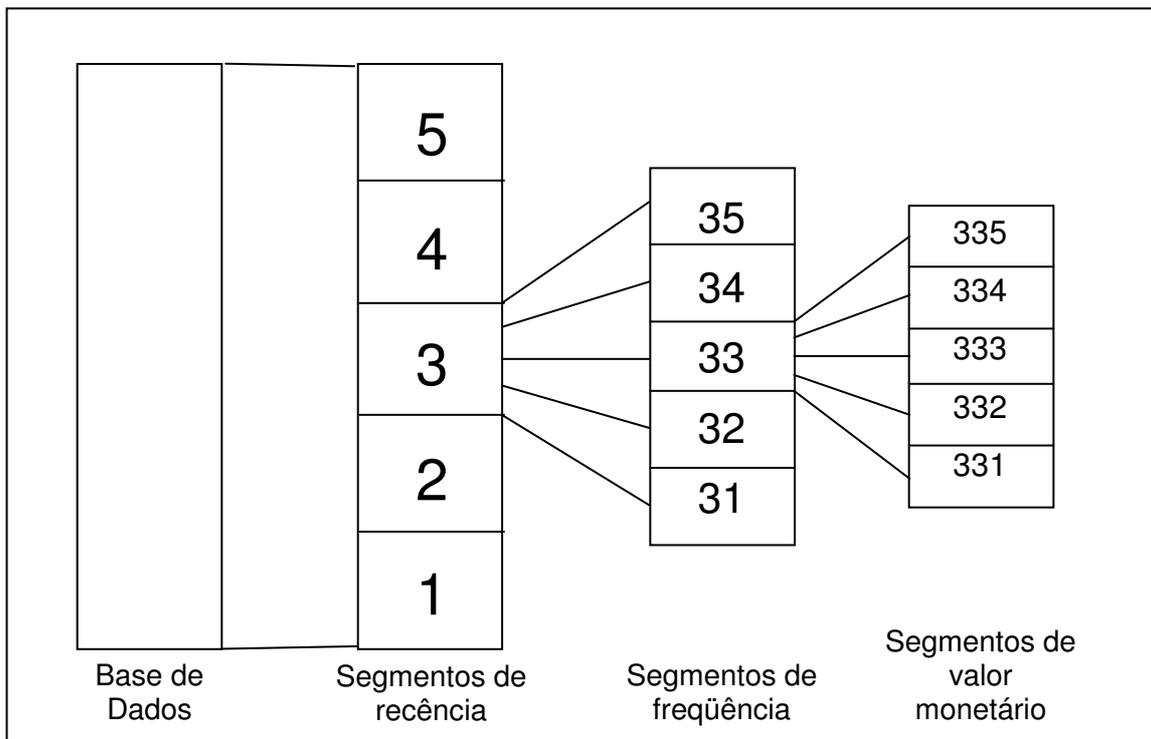


Figura VI: Ordenação e divisão das células RFV

Nesta seção será apresentado um exemplo prático de Análise RFV utilizando-se uma base fictícia com um número reduzido de registros de clientes. O objetivo deste exemplo é ilustrar os conceitos de segmentação de clientes em células de RFV e suas aplicações. Suponha os seguintes dados fictícios apresentados na Tabela I.

A primeira coluna da tabela (Id) é um identificador único seqüencial do registro de cliente. As demais colunas possuem as informações necessárias à classificação RFV, que são a quantidade de dias passados desde a última compra realizada pelo cliente (Última Compra), a quantidade de transações realizadas pelo cliente no período considerado (Transações) e o valor total gasto pelo cliente neste mesmo período (Total Gasto).

Para que seja feita a classificação pela dimensão de recência deve-se ordenar os clientes de forma decrescente, do mais recente para o menos recente. No exemplo, em que a quantidade de dias é fornecida, a ordenação fica conforme apresentada na Tabela II.

Id	Última Compra	Transações	Total Gasto
	(dias)	(qtd)	(R\$)
1	100	10	426
2	69	3	83
3	42	6	231
4	104	6	491
5	107	5	340
6	3	3	307
7	75	9	252
8	27	3	161
9	70	8	136
10	45	7	215
11	45	4	336
12	9	10	113
13	30	9	162
14	61	8	399
15	131	7	173
16	23	3	450
17	71	6	408
18	57	8	413
19	144	6	72
20	146	7	114

Tabela I: Dados fictícios de clientes para classificação RFV

Deve-se, então, dividir os clientes em um determinado número de células de recência. Neste exemplo decidiu-se criar cinco segmentos de recência. Cada uma destas células, de tamanho igual, recebe uma numeração que é o índice da recência. Ver Tabela III.

Para classificar esta base de clientes na dimensão de freqüência deve-se fazer a ordenação dos registros de cada célula de recência pelos dados de freqüência. Assim como na recência, a ordenação deve ser dos clientes mais freqüentes para os menos freqüentes. Neste exemplo escolhemos, arbitrariamente, dividir as células de freqüência em somente dois grupos. Este procedimento está exemplificado na Tabela IV.

Observe que o cliente com Id 12 recebe os índices de recência 5 e freqüência 2. Dividindo-se a base de clientes em cinco segmentos de recência e dois segmentos de freqüência obtém-se 10 células distintas ($5 \times 2 = 10$).

O próximo passo é dividir cada célula recência-freqüência em segmentos de valor monetário. Digamos que, assim como freqüência, serão criados dois segmentos de valor monetário. Com esta configuração, 20 diferentes grupos são formados ($5 \times 2 \times 2 = 20$). A Tabela V apresenta a classificação final.

Neste exemplo simplificado, cada um dos grupos RFV possui apenas um cliente, o que pode não ocorrer na realidade. O grupo 522, por exemplo, representa os clientes com maior classificação possível nas dimensões de recência, freqüência e valor. Neste caso o único cliente classificado no grupo em questão é o que possui Id 12

2.5. CONFIGURAÇÃO DA ANÁLISE RFV

Conforme pôde ser observado na seção anterior, para que a Análise RFV seja aplicada, certas decisões precisaram ser tomadas. Além de arbitrariamente determinar a quantidade de células que serão formadas pela combinação de cortes em cada uma das dimensões (R, F e V), é necessário definir que período deve ser considerado na análise e a ordem na qual as

diversas dimensões serão aplicadas na ordenação dos segmentos (RFV, FRV, VRF etc.).

Cada um dos domínios possíveis para estas variáveis afetará diretamente o resultado da previsão de retorno de cada grupo e, por conseqüência, terá influência no resultado estimado e real das campanhas de marketing. Selecionar um período muito curto, por exemplo, pode significar deixar muitos bons clientes de fora do público-alvo. Um período muito longo, por outro lado, fará com que clientes há muito inativos influenciem na classificação dos demais.

Da mesma forma, dividir a amostra em apenas um grupo de recência, colocará no mesmo segmento clientes muito distintos em termos de perfil de compras. Dividi-los em um número excessivo de segmentos, por outro lado, fará com que o número de casos em cada segmento torne-se muito pequeno. A medida que o número de casos em cada grupo diminui, o poder de previsão do resultado do segmento fica reduzido. Por exemplo, se um instituto de

Id	Última Compra	Recência	Transações	Total Gasto
	(dias)		(qtd)	(R\$)
6	3	5	3	307
12	9		10	113
16	23		3	450
8	27		3	161
13	30	4	9	162
3	42		6	231
10	45		7	215
11	45		4	336
18	57	3	8	413
14	61		8	399
2	69		3	83
9	70		8	136
17	71	2	6	408
7	75		9	252
1	100		10	426
4	104		6	491
5	107	1	5	340
15	131		7	173
19	144		6	72
20	146		7	114

Tabela III: Dados de clientes classificados de acordo com recência.

pesquisa eleitoral perguntar a intenção de voto a apenas uma ou duas pessoas, o grau de confiança da estimativa é praticamente nulo.

Bitran e Mondschein (1997, p.108) afirmam ainda que a relação inversa da recência com a taxa de resposta não é sempre verdadeira. Este caso é exemplificado para os bens duráveis como carros e imóveis. A relação pode, ainda, ser diferente dependendo do mercado, da oferta e da conjuntura econômica.

Fica claro, portanto, que existe um ponto ótimo entre as possíveis escolhas para cada uma das variáveis citadas. Esta configuração ótima vai depender de um modelo aplicado caso a caso. Ainda segundo Schmid (1998, p.38), “cada base de consumidores de catálogo possui certos hábitos e padrões de compra”. Hughes (2000, p.122) afirma que “existem várias maneiras de se fazer a Análise RFV. Os usuários terão que experimentar para descobrir qual funciona melhor para sua situação”.

Id	Última Compra	Recência	Transações	Frequência	Total Gasto
	(dias)		(qtd)		(R\$)
12	9	5	10	2	307
6	3		3		113
16	23		3	1	450
8	27		3		161
13	30	4	9	2	162
10	45		7		231
3	42		6	1	215
11	45		4		336
18	57	3	8	2	413
14	61		8		399
9	70		8	1	83
2	69		3		136
1	100	2	10	2	408
7	75		9		252
17	71		6	1	426
4	104		6		491
15	131	1	7	2	340
20	146		7		173
19	144		6	1	72
5	107		5		114

Tabela IV: Dados de clientes classificados de acordo com recência e frequência.

O modelo proposto nesta tese determina o ponto ótimo para todas estas escolhas e é baseado em técnicas de inteligência computacional fazendo amplo uso de Algoritmos Genéticos, assunto abordado no próximo capítulo.

Id	Última Compra (dias)	Recência	Transações (qtd)	Frequência	Total Gasto (R\$)	Valor
12	9	5	10	2	307	2
6	3		3		113	1
16	23		3	1	450	2
8	27		3		161	1
10	45	4	7	2	231	2
13	30		9		162	1
11	45		4	1	336	2
3	42		6		215	1
18	57	3	8	2	413	2
14	61		8		399	1
2	69		3	1	136	2
9	70		8		83	1
1	100	2	10	2	408	2
7	75		9		252	1
4	104		6	1	491	2
17	71		6		426	1
15	131	1	7	2	340	2
20	146		7		173	1
5	107		5	1	114	2
19	144		6		72	1

Tabela V: Clientes classificados de acordo com recência, frequência e valor.

Capítulo 3. ALGORITMOS GENÉTICOS

3.1. INTRODUÇÃO

A ciência dos Algoritmos Genéticos surgiu nos anos das décadas de 1950 e 1960 e vem sendo continuamente desenvolvida ao longo do tempo. Atualmente utilizada nas mais diversas áreas como ferramenta de otimização, Algoritmos Genéticos também são utilizados em outros segmentos como geração automática de código, definição de carteiras de investimento, entre outras.

A escolha desta técnica para realizar a otimização dos resultados das campanhas de marketing direto, buscando uma configuração ótima da Análise RFV está baseada nos pontos apresentados por Goldberg (1989):

- Inúmeros casos de sucesso no uso do algoritmo têm sido relatados nas mais diversas áreas de conhecimento;
- É um algoritmo extremamente robusto, sendo capaz de localizar pontos de máximo e mínimo em espaços de busca de grande complexidade;
- O mecanismo de mutação embutido no algoritmo aumenta as chances de que o mecanismo de otimização não fique preso em um máximo ou mínimo local;
- Ampla disponibilidade de software que implementa o algoritmo com eficiência.

3.2. BREVE HISTÓRICO

Algoritmos Genéticos fazem parte de uma ciência mais abrangente chamada “Computação Evolucionária”, que surgiu nos anos 50s e 60s quando

vários cientistas computacionais, tais como Rechenberg (1965), Fogel *et.al* (1966), Box (1957), Friedman (1959), Bledsoe (1961), Bremermann (1962), Reed *et.al* (1967) e outros, de forma independente, se seduziram pela idéia de que as teorias evolutivas poderiam ser utilizadas como ferramentas de otimização para problemas de engenharia.

A idéia que imperava na época era que um conjunto inicial de soluções candidatas poderiam ser submetidas a um processo de evoluções contínuas, utilizando operadores inspirados na diversidade genética e na seleção natural, até que atingissem soluções ótimas. O primeiro resultado concreto, em termos de aplicações práticas, foi obtido por Rechenberg (*op. cit.*), nos anos 60s, através da elaboração de um método utilizado para otimizar, com sucesso, estruturas semelhantes a aerofólios chamado “Estratégias Evolutivas” (do original em alemão *Evolutionstrategie*).

Um ano depois Fogel, Owens e Walsh (*op. cit.*), inspirados pelos trabalhos de Rechenberg (*op. cit.*), desenvolveram a chamada “Programação Evolutiva”, uma técnica onde soluções candidatas eram representadas como estados de uma máquina. Estes estados evoluíam de forma aleatória, através de mutações, buscando selecionar a configuração da máquina mais apta a resolver um determinado problema.

Os Algoritmos Genéticos, da forma como são conhecidos nos dias de hoje, foram concebidos e desenvolvidos por John Holland na Universidade de Michigan entre os anos 60s e 70s (HOLLAND, 1975). Curiosamente os objetivos de Holland não eram utilizar Algoritmos Genéticos como método de busca de soluções ótimas para problemas específicos, mas sim possibilitar que o estudo do fenômeno da adaptação, conforme ele ocorre na natureza, fosse realizado em um ambiente computacional.

O método de Algoritmo Genético proposto por Holland (*op. cit.*) inicia o processo de busca de uma solução ótima com uma população de

cromossomos, representados por *strings* de *bits*⁸, que evolui na direção de populações cada vez mais aptas. A evolução destas populações baseia-se em um processo de seleção natural com operadores inspirados na combinação de genes que ocorre naturalmente na reprodução dos seres vivos.

O método proposto por Holland (*op. cit.*) baseia-se na existência de três operadores genéticos, a saber: inversão, mutação e *crossover*⁹ (também referenciada na literatura como recombinação cruzada, recombinação genética, recombinação sexual ou apenas cruzamento).

Atualmente as implementações de Algoritmos Genéticos mais comuns seguem as mesmas premissas descritas por Holland (*op. cit.*), com exceção do operador de inversão cujos benefícios, se existirem, ainda não foram provados (MITCHELL, 1999).

A pesquisa e utilização das técnicas de Algoritmos Genéticos tem evoluído muito nos dias atuais. Em 1989 o primeiro sistema que fazia uso de Algoritmos Genéticos foi lançado comercialmente (PALISADE, 2001) e até março de 2003, um grupo de pesquisadores americanos de Los Altos, Califórnia, conseguiu “reinventar” quinze invenções humanas no ramo da eletrônica utilizando Algoritmos Genéticos. Seis delas haviam sido patenteadas recentemente por grandes institutos de pesquisa.

Uma das reinvenções apresentou uma melhoria em relação à invenção original que foi suficiente para permitir a solicitação de patente. Se o pedido de patente for aceito, será a primeira patente na história da humanidade concedida a uma invenção criada por programação genética (KOZA *et. al*, 2003).

Mitchell (1999) e Koza, Keane e Streeter (2003) fornecem mais detalhes sobre a história da computação evolutiva e a posição atual desta ciência.

3.3. COMO OS ALGORITMOS GENÉTICOS FUNCIONAM

⁸ Seqüência de zeros e uns.

⁹ Recombinação de genes através da troca genética entre cromossomos.

O principal conceito que está por trás dos Algoritmos Genéticos é a Teoria Evolutiva enunciada pela primeira vez por Charles Robert Darwin no século XIX (DARWIN, 1859). Segundo Darwin (*op. cit*), a Teoria Evolutiva está baseada em três regras simples, a saber:

- i) A evolução ocorre no nível dos cromossomos através da reorganização de genes.
- ii) A natureza tende a permitir mais reproduções dos organismos com maior capacidade de adaptação.
- iii) A diversidade é mantida na população graças a mutações aleatórias.

3.3.1. UM EXEMPLO DARWINIANO

Observe que, em geral, estas regras levam a uma população com seres cada vez mais adaptados ao seu ambiente. Por exemplo, imagine que a população observada se trata de um conjunto de ratos que podem ser classificados segundo dois critérios: tamanho e coloração. Quanto ao tamanho os ratos podem ser grandes ou pequenos e sua coloração pode ser clara ou escura. Esta população está distribuída conforme a figura VII.

Com o passar do tempo, alguns gatos mudam-se para o mesmo ambiente. Os gatos, como predadores de ratos, constituem uma alteração importante no ambiente. Conforme os gatos vão caçando os ratos percebe-se uma mudança no perfil da população na medida que novas gerações são formadas.

Os ratos menores e escuros são capturados com mais dificuldade e, por isso, acabam sobrevivendo por mais tempo e se reproduzindo. A reprodução destes ratos geram novas gerações com características herdadas de seus pais. Simplificando, ratos escuros e pequenos se cruzam e geram novos ratos também escuros e pequenos.

A diversidade na população é mantida porque os ratos pequenos, porém claros, também são menos capturados que os ratos maiores. Alguns ratos

maiores ainda conseguem sobreviver tempo suficiente para reproduzir, ainda que em menor quantidade. Eventualmente uma classe de ratos (grandes e claros) pode ser extinta desta população.

Ao longo de várias gerações uma mutação pode ocorrer. Esta mutação poderia, digamos, dar origem a ratos pequenos e malhados. Estes ratos poderiam ser menos vulneráveis ao predador que os ratos pequenos e claros e, portanto, aumentar em número durante as novas gerações.

Ao longo do tempo a população observada poderia possuir o aspecto apresentado na figura VIII. Note que, em média, esta é uma população mais adaptada ao ambiente (com gatos predadores) do que a população inicial.

3.3.2. ALGORITMOS GENÉTICOS COMPUTACIONAIS

Para a comunidade de pesquisadores, usuários e implementadores de Algoritmos Genéticos, os mecanismos da evolução natural descritos por Darwin (1859) são adequados não somente para explicar a evolução das espécies, mas também para a solução de uma gama variada de problemas nas mais diversas áreas de conhecimento, especialmente aqueles que pressupõem a busca de soluções ótimas em um espaço relativamente grande de soluções possíveis (MITCHELL, 1999).

Neste sentido os Algoritmos Genéticos implementam as três regras estabelecidas por Darwin (*op. cit*) através de um processo iterativo de transformação de um conjunto de *strings* de *bits*, onde cada *string* representa

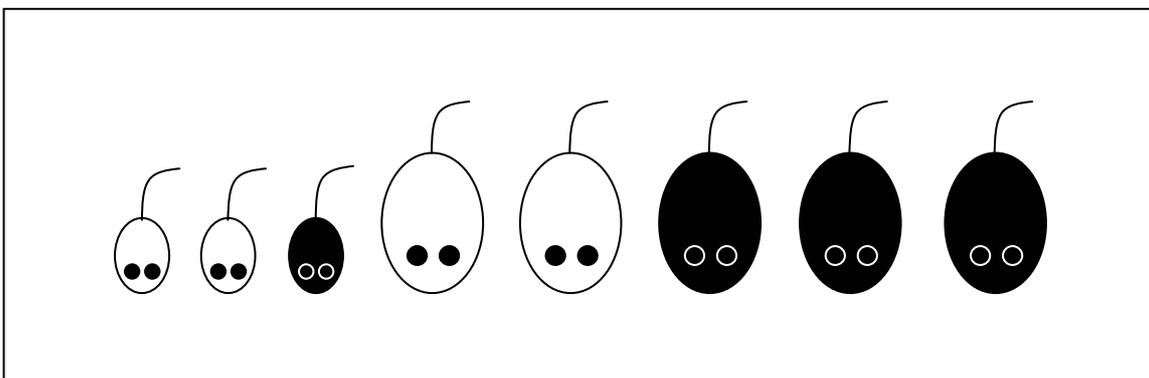


Figura VII: População de ratos.

uma possível solução para um determinado problema que se deseja resolver. Por exemplo, na determinação da localização ideal dos itens de estoque em um depósito, cada *string* representaria uma possível organização física destes itens nos diversos setores, estantes e prateleiras do depósito em questão. Na determinação de uma escala ótima de pilotos em uma companhia aérea, cada *string* representa uma escala possível de ser utilizada pela companhia em questão.

Na nomenclatura padrão de Algoritmos Genéticos o conjunto de *strings* é chamado de população e as *strings* propriamente ditas de cromossomos. Cada cromossomo é dito representar um organismo.

Sobre a população podem ser executadas quatro operações matemáticas, a saber: seleção, *crossover*, mutação e inversão. Estas operações procuram fazer com que os organismos, representados pelos cromossomos, se tornem cada vez mais aptos à sobreviverem no seu ecossistema. Em outras palavras, transforma as *strings* em soluções cada vez melhores para o problema que se deseja resolver.

O operador de seleção elege os pares de cromossomos (*string* de *bits*) da população que estão mais aptos a se reproduzirem. O operador de *crossover* produz um novo par de cromossomos, chamados de *offsprings*¹⁰, com subpartes dos pares de cromossomos reprodutores, imitando o processo de

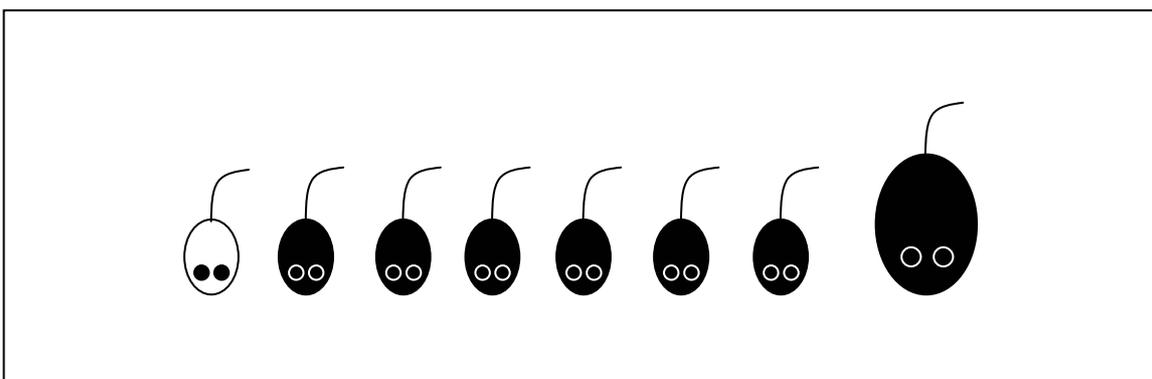


Figura VIII: População de ratos em ambiente com novos predadores.

¹⁰ Filhos ou filhotes.

combinação biológica entre dois organismos. O operador de mutação altera aleatoriamente partes dos *offsprings* que foram gerados pelo operador de *crossover*. Finalmente o operador de inversão inverte toda uma cadeia de *bits* de cada *offspring*. Ver Figura IX.

Enquanto o operador de *crossover* sempre produz pares de *offsprings* a partir dos pares de cromossomos reprodutores, os operadores de mutação e inversão raramente efetuam alguma modificação sobre os *offsprings*, uma vez que na natureza as mutações aleatórias que estes operadores procuram simular ocorrem com baixíssima frequência. É importante mencionar que nas modernas implementações de Algoritmos Genéticos o operador de inversão não é utilizado, visto que seus benefícios, se é que existem, não foram ainda comprovados (MITCHELL, 1999, p. 8).

Obviamente o processo iterativo utilizado pelos Algoritmos Genéticos não é

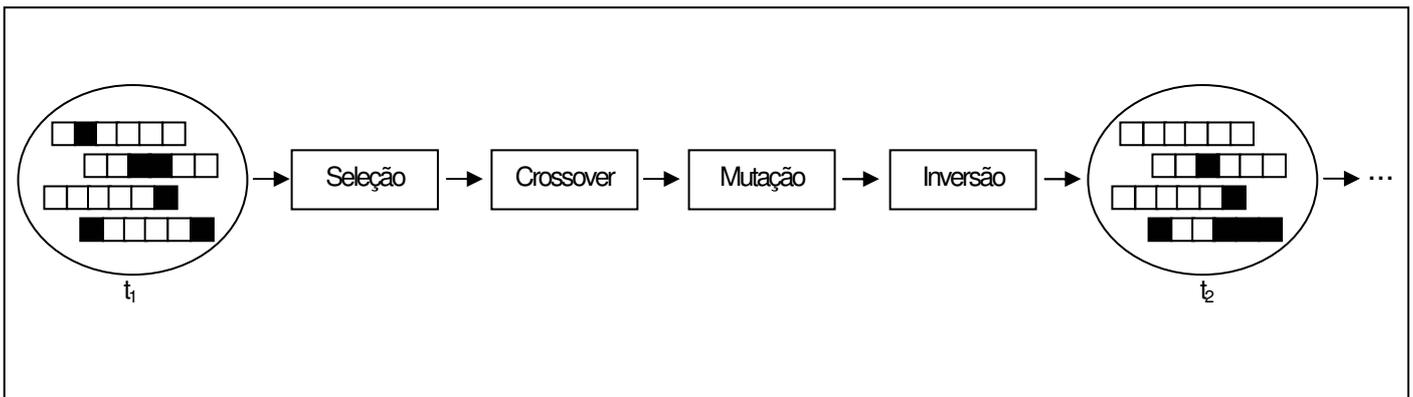


Figura IX: Operadores utilizados no processo

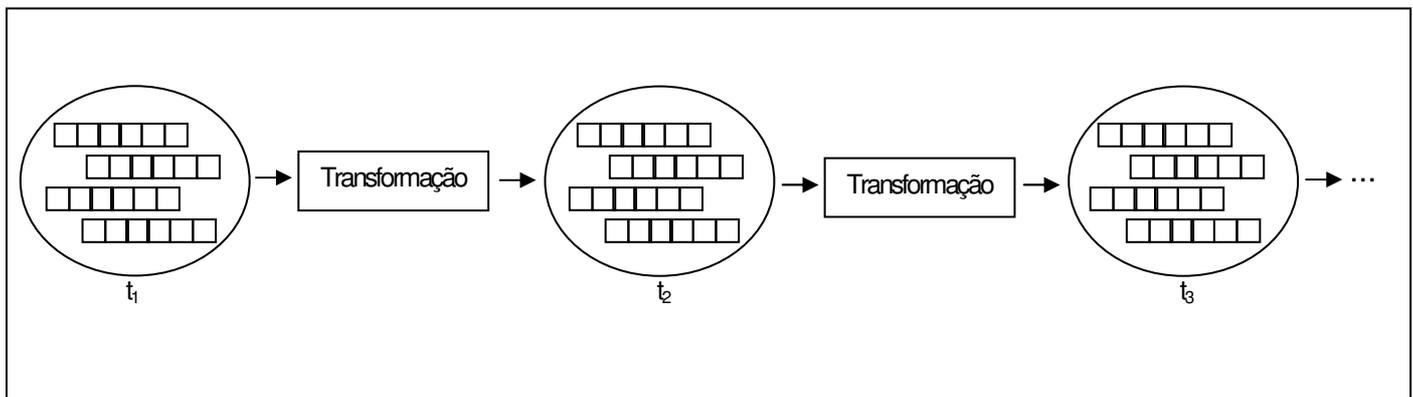


Figura X: Processo de transformação da população.

executado indefinidamente. Cada cromossomo que faz parte da população em um determinado instante é avaliado através de uma função de aptidão (*fitness function*), que determina quão apto a sobreviver no seu ecossistema o cromossomo está, isto é, quão próximo cada *string* está da solução ideal do problema que se deseja resolver.

O processo termina quando uma ou mais das seguintes condições forem atingidas: a solução ideal for encontrada, já se está suficientemente próximo da solução ideal ou um determinado número de tentativas para se obter a solução ideal foi atingido.

3.3.3. DETALHANDO O OPERADOR DE SELEÇÃO

O operador de seleção determina quais organismos serão selecionados para a reprodução e geração de *offsprings* através da operação de *crossover*. Quanto mais apto a sobreviver for o organismo, maior é a probabilidade dele ser selecionado para a reprodução. A aptidão de cada organismo é estabelecida por uma função de aptidão, ou seja, *fitness function*, que é escolhida pelo usuário de Algoritmos Genéticos baseada na natureza do problema. Ver item 3.2.5.

Algumas implementações de Algoritmos Genéticos utilizam este operador não somente para determinar quais organismos participarão de processos de reprodução, mas também que organismos serão mantidos na próxima geração através de um processo de cópia (KOZA *et. al*, 2003).

Guervós (2003) apresenta três diferentes formas de se aplicar a seleção em um processo de otimização baseado em Algoritmos Genéticos:

- Baseado em *Ranking* - neste esquema uma quantidade arbitrária da população é mantida inalterada na geração seguinte. Os “n” organismos menos aptos são eliminados e substituídos, através da operação de *crossover*, por *offsprings* dos organismos restantes.

- Roda de Roleta - cria-se um conjunto arbitrariamente grande de organismos da geração atual, onde o número de ocorrências de um determinado organismo é proporcional ao seu grau de aptidão. Por exemplo, se um organismo A possui um grau de aptidão duas vezes superior ao do organismo B, o número de ocorrências de organismos A no conjunto em questão será o dobro do número de ocorrências de organismos B. Seleciona-se aleatoriamente neste conjunto os pares de organismos que irão se reproduzir. Observe que esta forma de implementação da operação de seleção permite que pares de organismos idênticos sejam selecionados para reprodução.
- Seleção por Torneio - seleciona-se aleatoriamente um número arbitrário de organismos. Aqueles que possuírem um maior grau de aptidão terão maior probabilidade de serem selecionados para reprodução.

Os diferentes processos de seleção garantem que os organismos mais aptos tenham uma maior probabilidade de ser selecionados para reprodução, gerando populações de elementos cada vez mais aptos a sobreviverem no seu ecossistema, isto é, cada vez mais próximos da solução ótima do problema que se deseja resolver.

3.3.4. DETALHANDO O OPERADOR DE CROSSOVER

A reprodução de organismos e geração de *offsprings* é realizada através do operador de *crossover*. Este operador realiza a combinação de sub-partes dos organismos para formar organismos novos com características herdadas dos organismos originais. Mitchell (1999) afirma que o cruzamento de organismos reprodutores com alto grau de aptidão pode otimizar o aparecimento de *offsprings* ainda mais aptos.

A Figura XI apresenta dois organismos selecionados para a reprodução que, combinados através do *crossover*, geram outros dois organismos inteiramente novos. Embora as características existentes nos organismos

reprodutores permanecerem nos *offsprings*, a recombinação de genes poderá gerar organismos com maiores ou menores graus de aptidão.

A operação de *crossover* apresentada na figura em questão pode ser chamada de *crossover* de ponto único, já que a troca de sub-partes entre os organismos reprodutores na geração de *offsprings* ocorreu em apenas um ponto de recombinação. Mas o *crossover* também pode ser realizado em múltiplos pontos, sendo a escolha de um ou dois pontos de recombinação as mais comuns (GUERVÓS, 2003, p.6).

A seleção dos pontos de recombinação é aleatória. Isto significa que em uma população onde os organismos possuem “n” *bits* ($n \geq 1$), $n - 1$ pontos possíveis podem ser selecionados e a reprodução de dois organismos pode gerar até $n - 1$ diferentes pares de *offsprings*.

Guervós (*Op. cit*) aponta que em alguns casos aplicar aleatoriamente o *crossover* pode gerar soluções inválidas para o problema que se quer otimizar. Nestes casos é necessário aplicar-se um *crossover* especializado que possua a inteligência necessária para evitar combinações incoerentes.

Por exemplo, se Algoritmo Genético está sendo utilizado para determinar a melhor escala possível de pilotos, em termos de custos operacionais, para uma companhia aérea, e a legislação proíbe que um piloto permaneça mais que um

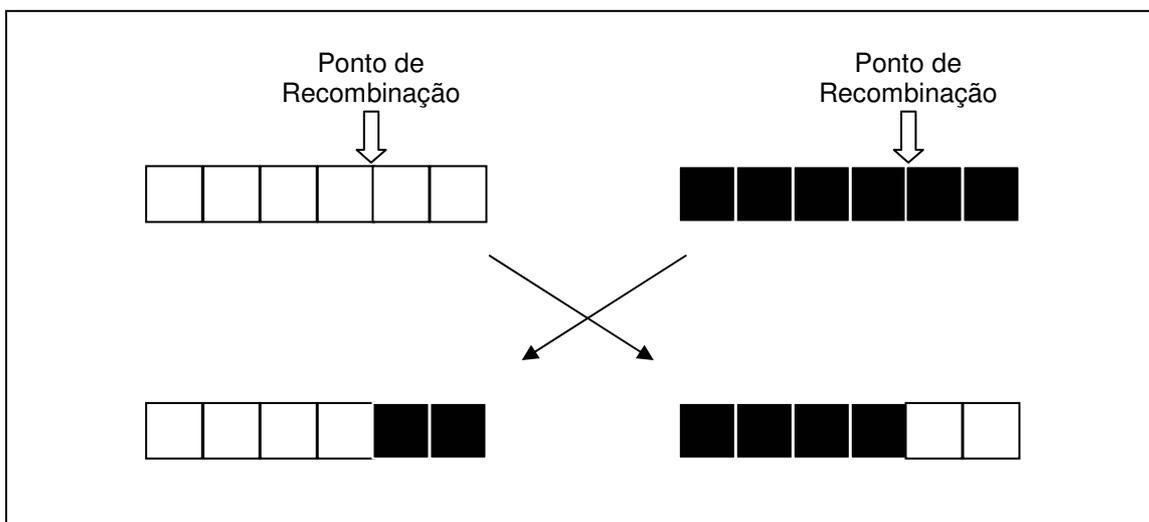


Figura XI: *Crossover* entre organismos

número determinado de horas trabalhando. Operações de *crossover* que gerem escalas onde a legislação é contrariada devem ser desconsideradas, mesmo que produzam custos operacionais menores.

Mitchell (1999, p.30) afirma que os pesquisadores da área acreditam que o operador de *crossover* é o principal responsável pelo poder de otimização dos Algoritmos Genéticos.

3.3.5. DETALHANDO O OPERADOR DE MUTAÇÃO

O operador de mutação é utilizado para manter a diversidade em uma população de organismos gerados a partir do processo de *crossover*. No processo de evolução natural dos organismos, a mutação com sucesso de sobrevivência é extremamente rara. Na maioria dos casos as mutações são letais, apesar disso é inegável que elas contribuem para a diversidade genética da espécie. Em Algoritmos Genéticos este operador tende ao mesmo papel, com a mesma frequência (GUERVÓS, 2003, p.7).

A mutação ocorre no nível do gene dos *offsprings* com uma probabilidade muito pequena. Cada gene possui a mesma probabilidade de sofrer uma mutação que, caso ocorra, modifica aquele gene em particular no organismo mutante.

Por exemplo, o organismo representado pela *string* "0100101" em um processo de otimização baseado em Algoritmos Genéticos poderia sofrer uma mutação, se transformando no organismo "0000101". Pode acontecer de um organismo sofrer mutação em mais de um *bit*, mas dado que a probabilidade da mutação ocorrer é muito pequena, este é um fenômeno extremamente raro.

Guervós (*Op. cit.*) alerta, ainda, que não deve-se abusar do operador de mutação, pois apesar de ser um mecanismo importante para manter a diversidade, seu uso intensivo pode transformar o Algoritmo Genético em uma busca aleatória. Outros métodos possíveis para aumentar a diversidade de uma população de cromossomos são o simples aumento do tamanho da

população e a garantia de que a população inicial é gerada aleatoriamente. Na Figura XII a ocorrência da mutação após o *crossover* é apresentada na forma de um diagrama.

3.3.6. DETALHANDO O OPERADOR DE INVERSÃO

O operador de Inversão atua, também, na manutenção da diversidade da população. Utilizando-se uma probabilidade, geralmente muito pequena, determina-se se um organismo gerado a partir de uma reprodução (*crossover*) sofrerá ou não a inversão, que consiste na troca de todos os *bits* do organismo.

Se, por exemplo, o organismo representado pela string “010001” sofrer uma inversão, o organismo gerado possuirá a seguinte representação: “101110”. Note que cada *bit* do organismo original foi invertido, gerando um novo organismo totalmente diferente.

3.3.7. DETALHANDO A FUNÇÃO DE APTIDÃO

A função de aptidão determina quão apto um organismo da população está a sobreviver e reproduzir. O resultado desta função serve de base para definir

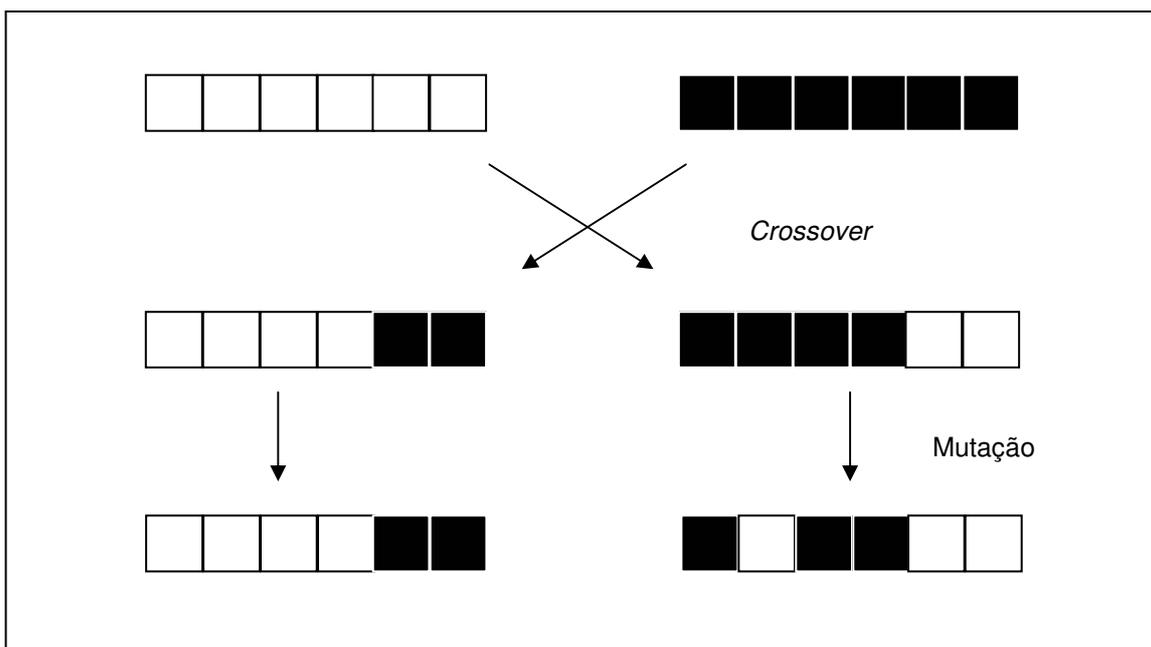


Figura XII: *Crossover* entre organismos. Um dos *offsprings* sofre mutação.

a probabilidade que um organismo terá de participar de uma reprodução e transportar os seus genes para a próxima geração. É este processo de seleção e reprodução, somado à mutação, que garante que a população de organismos estará evoluindo na direção de soluções mais otimizadas.

Por exemplo, se um Algoritmo Genético está sendo utilizado para buscar uma solução otimizada para o tradicional problema do caixeiro viajante, que deve visitar um determinado número de cidades com distâncias variadas percorrendo o menor caminho possível, pode-se dizer que uma função de aptidão adequada é o inverso da distância total percorrida. Quando menor a distância percorrida pelo caixeiro em uma determinada solução, mais apta esta solução estará a sobreviver e reproduzir.

3.3.8. EXEMPLO COMPUTACIONAL

Para efeitos de entendimento, apresentamos um exemplo da aplicação de Algoritmos Genéticos em um sistema computacional. Todas as operações e funções apresentadas a seguir foram escolhidas de forma a facilitar o entendimento de uma aplicação de Algoritmo Genético em um ambiente computacional. Considere, portanto, a função $g(x) = -x^2$ para $x \in \mathbb{Z}$, onde \mathbb{Z} é o conjunto dos inteiros. A Figura XIII apresenta o gráfico de $g(x)$ para $x \in [-15,15]$.

Suponha que estamos interessados em obter o máximo de $g(x)$, utilizando

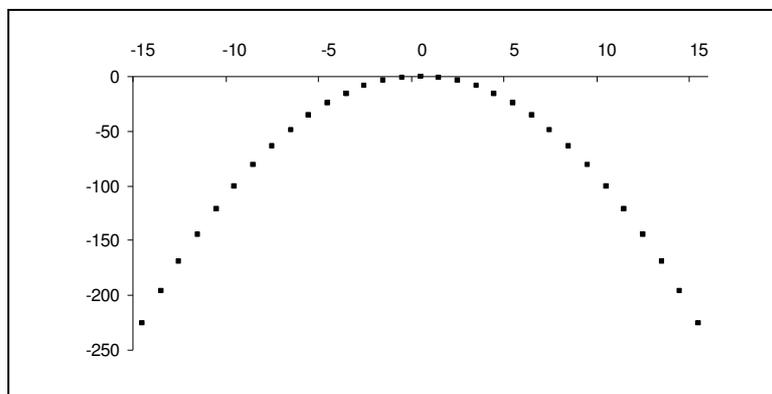


Figura XIII: Representação Gráfica da função $g(x)$.

Algoritmos Genéticos. Para representar os organismos que definem as soluções possíveis para o problema, serão utilizadas *strings* de cinco *bits*. Neste contexto as *strings* “10010”, “00000” e “01010” são representações válidas de possíveis soluções para o problema. Isso nos possibilita avaliar até 32 soluções diferentes para a obtenção do máximo de $g(x)$, visto que $2^5 = 32$.

Na nomenclatura que iremos adotar cada *bit* é representado pela letra “b” e os *bits* são numerados da direita para a esquerda. O *bit* mais à direita é o *bit* 0 e o *bit* mais à esquerda o *bit* 4. Neste caso, os *bits* $b_4 = 1, b_3 = 0, b_2 = 0, b_1 = 1$ e $b_0 = 0$ formam a *string* “10010”. Por simplicidade representamos uma *string* qualquer por “ $b_4 b_3 b_2 b_1 b_0$ ”.

Dado que cada *string* representa um possível valor no conjunto dos inteiros, existe uma função $h: S \rightarrow Z$, onde S é o conjunto das *strings* de cinco *bits*, tal que:

$$h(\text{“}b_4 b_3 b_2 b_1 b_0\text{”}) = k \times (2^3 b_3 + 2^2 b_2 + 2^1 b_1 + 2^0 b_0), \text{ e}$$

$b_4 = 0 \Rightarrow k = 1$ e $b_4 = 1 \Rightarrow k = -1$. Observe que o *bit* b_4 é o *bit* do sinal, que pode ser positivo ou negativo. Por exemplo,

$$h(\text{“}10010\text{”}) = -1 \times (2^3 \times 0 + 2^2 \times 0 + 2^1 \times 1 + 2^0 \times 0) = -2$$

e

$$h(\text{“}00010\text{”}) = 1 \times (2^3 \times 0 + 2^2 \times 0 + 2^1 \times 1 + 2^0 \times 0) = 2.$$

Esta foi a função de transformação escolhida para desenvolver o exemplo computacional nesta tese. Existem, entretanto, muitas outras possibilidades de fórmulas. A escolha da fórmula que gera o significado de uma *string* de *bits* em uma implementação de Algoritmos Genéticos faz parte da implementação do algoritmo e não deve influenciar na capacidade de otimização do mesmo.

É importante observar que na representação escolhida para *strings*, dois organismos aparentemente distintos correspondem ao mesmo inteiro, isto é, $h(\text{“}10000\text{”}) = -0 = 0 = +0 = h(\text{“}00000\text{”})$. Para evitar ambigüidade de

representação, escolhamos a *string* “00000” para representar o valor 0. Em decorrência, sempre que a *string* “10000” surgir em uma população, durante o processo de otimização, ela é imediatamente substituída pela *string* “00000”.

Evitar que *strings* indesejáveis surjam durante o processo de otimização é um procedimento comumente utilizado em Algoritmos Genéticos, como nos mostra Mitchell (1999).

Antes que possamos aplicar Algoritmos Genéticos para calcular o máximo de $g(x)$, precisamos definir a função de aptidão $f: P \rightarrow [0\%,100\%]$, onde P é uma população de organismos, que será utilizada para fornecer a probabilidade de seleção dos organismos que participarão das operações de *crossover*, gerando *offsprings*. A escolha da função de *fitness* vai variar de implementação para implementação, sendo portanto uma decisão de construção do modelo.

O importante é que a função de *fitness* reflita o objetivo que se deseja alcançar na otimização, ou seja, em outras palavras a função de *fitness* deve refletir quão bem um cromossomo resolve o problema a ser otimizado (MITCHELL, 1999, p. 9).

Neste caso desenvolvemos uma função de aptidão que reflete corretamente o objetivo que se deseja alcançar no exemplo:

$$f(p) = \frac{g(h(p)) - (\min(h(P)) + 1)}{\sum \{g(h(k)) - \min(P) - 1 \mid k \in P\}},$$

onde $\min(P) = \{p \mid p \in P \wedge \forall k \in P \Rightarrow h(k) \geq h(p)\}$. Por exemplo, seja

$$P = \{“01100”, “01010”, “10010”, “10100”, “10110”, “11111”\}.$$

Existem inúmeras outras funções de aptidão possíveis que gerariam o mesmo resultado.

Neste caso

	b_4	b_3	b_2	b_1	b_0	$h(p)$	$g(h(p))$	$g(h(p)) - (\min(P) + 1)$	$f(p)$	
P {	1	1	0	0	1	0	-2	-4	222	26,7%
	2	1	0	1	0	0	-4	-16	210	25,3%
	3	1	0	1	1	0	-6	-36	190	22,9%
	4	0	1	0	1	0	10	-100	126	15,2%
	5	0	1	1	0	0	12	-144	82	9,9%
	6	1	1	1	1	1	-15	-225	1	0,1%
Total								831	100%	

Tabela VI: População inicial e resultado das funções.

O cálculo realizado para obter-se o resultado da função de aptidão do primeiro organismo foi:

- (i) Organismo "10010", $h(p) = -1 \times (2^3 \times 0 + 2^2 \times 0 + 2^1 \times 1 + 2^0 \times 0) = -2$;
- (ii) $g(h(p)) = g(-2) = -2^2 = -4$;
- (iii) $g(h(p)) - (\min(p) - 1) = -4 - (\min(p) - 1) = -4 - (-225 - 1) = -4 + 226 = 222$;

Observe que $\sum \{g(h(k)) - \min(P) - 1 \mid k \in P\} = 831$, que $\sum f(p) = 100\%$ e que os organismos que possuem o maior *fitness* são os que geram maiores resultados para a função que desejamos maximizar, isto é $g(x)$. Estes valores estão condensados na tabela VII, que representa a população inicial, isto é, a população no instante t_0 , que iremos utilizar para calcular o máximo de $g(x)$.

Dentre os organismos que compõem a população inicial, o que está mais próximo da solução ótima é o organismo 1 e o que está mais distante é o organismo 6. Dado que a função de *fitness* indica a probabilidade de um organismo ser selecionado para participar do processo de reprodução, é de se esperar que o organismo 1 tenha uma probabilidade de ser selecionado maior que o organismo 6. Neste caso, a probabilidade do organismo 1 ser selecionado é 26,7% e a probabilidade do organismo 6 é apenas 0,1%.

	Organismos					g(h(s))	Fitness
1	1	0	0	1	0	-4	26,7%
2	1	0	1	0	0	-16	25,3%
3	1	0	1	1	0	-36	22,9%
4	0	1	0	1	0	-100	15,2%
5	0	1	1	0	0	-144	9,9%
6	1	1	1	1	1	-225	0,1%

Tabela VII: Organismos na população em t_0 .

A nova população gerada pelo processo de otimização, isto é a população no instante t_1 , está representada na tabela VIII.

	Organismos					Res.	Fitness
1	1	0	0	1	0	-4	22,8%
2	1	0	0	1	0	-4	22,8%
3	0	0	0	0	0	0	23,7%
4	1	0	1	1	0	-36	15,3%
5	1	1	0	1	0	-100	0,2%
6	0	0	1	1	0	-36	15,3%

Tabela VIII: Organismos da população em t_1 .

Para formar esta nova população, as seguintes operações ocorreram:

- O organismo 1 em t_0 foi selecionado para *crossover* com ele mesmo. Nesta implementação de Algoritmos Genéticos é possível selecionar o mesmo organismo para compor o par de cromossomos necessário ao *crossover*.

Desta forma, independentemente de onde o ponto de recombinação foi selecionado, os *offsprings* gerados pelo organismo 1 são cópias dele mesmo em t_0 . Estes *offsprings* são os organismos 1 e 2 de t_1 ;

- A segunda operação de *crossover* foi realizada entre os organismos 1 e 2, com o ponto de recombinação entre os *bits* b_2 e b_1 . A operação de *crossover* deu origem aos organismos representados pelas *strings* “10000” e “10110”. O organismo “10000” foi substituído pelo “00000” para evitar duplicidade de representações (+0 e -0). Estes são os organismos 3 e 4 de t_1 ;
- A terceira operação de *crossover* ocorreu entre os organismos 2 e 4, com o ponto de recombinação entre os *bits* b_4 e b_3 . Esta operação deu origem aos organismos 5 e 6 de t_1 , sendo que o organismo 6 sofreu uma mutação que inverteu o valor do *bit* b_4 .

Observe que a população no instante t_1 apresenta organismos mais próximos da solução do problema proposto do que no instante t_0 , isto é, em média possuem valores mais próximos ao máximo de $g(x)$. Isso pode ser constatado pela variação da média de $g(x)$, -87,5 em t_0 e -30 em t_1 .

No instante t_1 o organismo que possui o maior *fitness* é o 3, com $g(h(s)) = 0$. Em termos de aptidão este organismo supera o organismo mais apto do instante t_0 , que tinha $g(h(s)) = -4$. De fato, o organismo 3 é o melhor que poderia surgir do processo de otimização, pois ele fornece o maior valor possível para $g(x)$.

Novos organismos continuam sendo formados a partir da reprodução dos organismos em t_1 . Sem perda de generalidade deixaremos de representar as populações destes organismos nos instantes $t_2, t_3 \dots t_n$.

A população em t_{n+1} é apresentada na Tabela IX. Observe que nesta população a *string* “00000” repete-se nos organismos 2, 3 e 4. Isso significa que os organismos com as características que os fazem ser os mais aptos vêm sobrevivendo e reproduzindo-se entre si. Neste exemplo, cada organismo com

esta característica possui uma probabilidade de 20,4% de sobrevivência. Em conjunto a probabilidade destes organismos serem selecionados para a reprodução é de mais de 60%, bem maior do que o organismo menos apto da população, isto é, o organismo 6, cuja a probabilidade de ser selecionado é de apenas 0,6%. Vale a pena mencionar que “00000” fornece o máximo para $g(x)$.

O critério de parada aplicado a esta implementação de Algoritmos Genéticos é o de gerar novas populações até que todos os seus elementos possuam as mesmas características. Em conseqüência, no instante t_m , onde $m > n+1$, a população assume o aspecto apresentado na Tabela X. Observe que no instante t_m todos os organismos fornecem o valor máximo para $g(x) = -x^2$.

		Organismos					Res.	<i>Fitness</i>
1	1	0	0	1	0	-4	18,2%	
2	0	0	0	0	0	0	20,4%	
3	0	0	0	0	0	0	20,4%	
4	0	0	0	0	0	0	20,4%	
5	0	0	0	0	1	-1	19,9%	
6	0	0	1	1	0	-36	0,6%	

Tabela IX: Organismos da população em t_{n+1} .

		Organismos					Res.	<i>Fitness</i>
1	0	0	0	0	0	0	16,7%	
2	0	0	0	0	0	0	16,7%	
3	0	0	0	0	0	0	16,7%	
4	0	0	0	0	0	0	16,7%	
5	0	0	0	0	0	0	16,7%	
6	0	0	0	0	0	0	16,7%	

Tabela X: Organismos da população no instante t_m .

Outros critérios de parada que são aplicados comumente em Algoritmos Genéticos são:

- o número de iterações, isto é, o processo de otimização para em t_n , onde n é um número arbitrado pelo usuário; e
- a variação média da função $g(h(s))$ entre dois instantes de tempo, isto é, $|\mu_{g(h(s))} \text{ em } t_i - \mu_{g(h(s))} \text{ em } t_j| \leq \xi$, onde $i - j$ é arbitrariamente escolhido pelo usuário.

3.4. POSSÍVEIS APLICAÇÕES DOS ALGORITMOS GENÉTICOS

Algoritmos Genéticos têm sido utilizados em um grande número de problemas nas mais diversas áreas de conhecimento. Abaixo são apresentados alguns exemplos destas aplicações.

3.4.1. OTIMIZAÇÃO

Koza, Keane e Streeter (2003) apresentam 15 circuitos eletrônicos otimizados com Algoritmo Genético, entre eles o circuito balun, que reduz a impedância em um circuito de baixa voltagem, patenteado em 2001.

Grosche *et al* (2001) apresentam uma abordagem eficiente do problema de otimizar agendas de vôo nas companhias aéreas utilizando Algoritmos Genéticos. Segundo os autores a abordagem tradicional utilizada para o problema é decompô-lo em sub-partes, submetendo cada uma delas a um processo de otimização linear. A otimização com Algoritmos Genéticos revelou-se mais eficiente por tratar o problema de forma completa, oferecendo também a possibilidade de explorar a otimização de certas sub-partes em paralelo.

Goldberg (1989) apresenta um estudo detalhado de como Algoritmos Genéticos podem ser utilizados nas áreas de otimização e inteligência computacional.

3.4.2. GERAÇÃO AUTOMÁTICA DE PROGRAMAS

A geração automática de código de programas de computador através do uso de Algoritmo Genético é explorada por Mitchell (1999, p.36) que oferece como exemplo um programa desenvolvido em Lisp que calcula o período orbital de um planeta.

Koza *et al* (2001) utiliza uma técnica chamada *Genetic Programming*¹¹, isto é, geração automática de programas de computador utilizando Algoritmos Genéticos, para desenvolver sistemas computacionais que solucionam problemas não triviais, tais como o desenvolvimento de circuitos eletrônicos de filtros de baixa frequência.

Geração automática de programas é explorada, em detalhes, por Banzhaf *et al* (1998).

3.4.3. INTELIGÊNCIA COMPUTACIONAL

Algoritmos Genéticos podem ser utilizados para apoiar tarefas de descobrimento de padrões realizadas por sistemas computacionais. Yang e Honavar (1998) defendem que problemas de classificação e descobrimento de padrões requerem uma seleção dos atributos que serão utilizados na classificação a partir de um conjunto maior de atributos disponíveis. Esta seleção, segundo os autores, pode ser realizada com sucesso se apoiada por Algoritmos Genéticos.

Outra aplicação possível é a utilização de Algoritmos Genéticos em conjunto com técnicas de Inteligência Computacional. Lewis *et al* (1992) utilizaram Algoritmos Genéticos para desenvolver um sistema de redes neurais que controla um robô com capacidades de locomoção.

¹¹ Programação Genética.

A combinação de Algoritmos Genéticos com outras técnicas de inteligência computacional, tais como lógica *fuzzy* e redes neurais, é explorada em detalhes por Jain *et al* (1998).

3.4.4. ECONOMIA

Bauer (1994) e Dawid (1999) mostram como Algoritmos Genéticos podem ser utilizados em ciências de base sociais como a economia. Ambos exemplificam várias aplicações de Algoritmos Genéticos no ramo de investimentos financeiros, em especial na gestão de *portfólios* de investimentos, aplicações em bolsa de valores e em *commodities*.

3.4.5. MARKETING

Kwon *et al* (2001) mostra como Algoritmos Genéticos podem ser utilizados com sucesso na personalização de e-mails. Eiben *et al* (1998) propõe um modelo de retenção de clientes baseado em Algoritmos Genéticos. Eiben *et al* (1996) compara as técnicas de Algoritmos Genéticos com técnicas mais tradicionais para otimização do resultado de campanhas de marketing direto.

3.4.6. OUTRAS ÁREAS DE CONHECIMENTO

Onwubolu (2002) e Gen *et al* (1996) utilizam técnicas de Algoritmos Genéticos para aumentar a eficiência de processos de produção e controle. Tzafestas (1997) propõe a gestão e uso de bases de conhecimento com auxílio de Algoritmos Genéticos. Davidor (1991) indica como Algoritmos Genéticos podem ser utilizados na construção de robôs mais inteligentes. Haupt *et al* (1998) utiliza Algoritmos Genéticos para aumentar a eficiência de *arrays* de antenas.

3.5. APLICAÇÃO DE ALGORITMOS GENÉTICOS À ANÁLISE RFV NA CONSTRUÇÃO DE UM MODELO DE OTIMIZAÇÃO

Esta tese se propõe a mostrar que, a despeito de todo o prestígio que a Análise RFV desfruta entre os profissionais de marketing, a aplicação de um modelo que faça uso de Algoritmos Genéticos pode melhorar substancialmente os resultados das campanhas de marketing direto.

A melhora deste resultado será obtida através da otimização das configurações da Análise RFV.

Para que o modelo construído seja robusto o suficiente para permitir a inclusão de novas dimensões, numéricas e categóricas, e seja capaz de lidar tanto com relações lineares e não lineares entre as variáveis, se faz necessário o uso de um algoritmo de otimização robusto e eficiente como os Algoritmos Genéticos.

Capítulo 4. FERRAMENTAS E MÉTODO DE PESQUISA

4.1. INTRODUÇÃO

Durante a construção deste modelo e desenvolvimento da tese, foram consideradas as hipóteses de que a Classificação RFV pode ser aplicada a diversos mercados com resultados positivos (HUGHES, 2000) mas que, porém, pode apresentar resultados diferentes de acordo com as decisões arbitrárias necessárias para configuração desta técnica.

Para otimizar este resultado-se propôs-se a utilização construção de um modelo que utiliza Algoritmos Genéticos. Para experimentar e apoiar a construção de um modelo foram utilizados dados reais de uma campanha de marketing direto realizado por uma das maiores cadeias de eletro-eletrônicos do país. Estes dados, bem como uma apresentação da empresa, são detalhados a seguir junto com as ferramentas de software utilizadas.

São apresentados:

- Os dados de uma campanha de marketing direto de uma grande cadeia de lojas de eletro-eletrônicos e móveis, a qual nos referenciaremos como “Empresa X”;
- O software Evolver for Excel 4, Industrial Version (WINSTON, 2000, PALISADE, 2001);
- O software Microsoft Excel 2000 (WALKENBACH, 1999);

4.2. A EMPRESA X

Fundada em 1946, no Rio de Janeiro, a Empresa X é uma das maiores e mais tradicionais redes de varejo de eletro-eletrônicos e móveis do país, estando presente nos principais estados brasileiros e mantendo-se, na última década, sempre como uma das cinco maiores empresas de seu mercado.

Em 1996, com o objetivo de acelerar o crescimento dos negócios, a Empresa X promoveu a abertura de seu capital com a emissão de 4,1 milhões de ações preferenciais na Bolsa de Valores de São Paulo, que foram complementadas pela emissão de *American Depositary Receipts*¹² (ADR) e US\$ 100 milhões de *Euronotes*¹³.

Nos dias de hoje, a empresa atua nacionalmente com sete centros de distribuição de mercadorias e mais de 350 lojas, possui mais de 8.000 empregados, faturou R\$ 2,7 bilhões em 2002 e obteve um lucro líquido de R\$ 35,8 milhões. O patrimônio líquido da Empresa X foi avaliado em R\$ 561 milhões em 2002. Neste mesmo ano a empresa efetuou mais de 10 milhões de transações com o ticket médio de cerca de R\$ 250,00.

Destaques Financeiros	2002	2001
Faturamento (R\$ milhões)	2.701	2.602
Número de Lojas	346	348
Número de Funcionários	8.370	7.409
Patrimônio Líquido (R\$ milhões)	561	547
Quantidade de Ações (x 1000)	89.758	63.765
Destaques Financeiros	2002	2001

¹² *American Depositary Receipts* são cotas de um investimento realizado em ações de uma empresa não americana, que são negociados na Bolsa de Valores de Nova Iorque (*New York Stock Exchange*). Os ADRs foram criados para permitir que ações negociadas em uma moeda estrangeira pudessem ser comercializadas em bolsas de valores americanas.

¹³ *Euronotes* são notas promissórias de curta duração lançadas no mercado europeu.

Lucro Líquido (R\$ milhões)	35,8	39,1
Lucro por Ação	0,40	0,61
Venda por m ² de loja (R\$ mil)	10,6	10,4
Vendas por Funcionário (R\$ mil)	322,7	351,2

Tabela XI: Destaques Financeiros da Empresa X

Todos estes números formam um conjunto impressionante de conquistas para uma empresa que começou a operar, há apenas 52 anos atrás, com uma única loja no centro do Rio de Janeiro e demonstram que este é um exemplo relevante para o desenvolvimento deste estudo.

Além do canal tradicional de venda direta ao consumidor através de suas lojas, a Empresa X realiza tanto vendas pela Internet (*e-business*) quanto por telefone (tele-vendas). As vendas pelo canal tradicional respondem por 95% do faturamento. Cerca de 70% das vendas realizadas pela Empresa X são pagas em parcelas (à prazo).

4.3. CAMPANHA UTILIZADA

Em Dezembro de 2002 a Empresa X resolveu contatar 50.529 pessoas de sua base de dados, com o objetivo de aumentar a circulação de pessoas dentro das lojas físicas.

A campanha em questão não efetuava a oferta de nenhum produto específico. Tratou-se, na verdade, de uma carta personalizada convidando os clientes, que já haviam efetuado compras anteriores, a retornarem a uma das lojas da rede e a efetuarem novas compras. Aos clientes que resolvessem aceitar a oferta, a Empresa X ofereceu um crédito pré-aprovado em sintonia com a faixa de renda de cada cliente.

A carta foi enviada pelo correio para clientes temporariamente inativos, isto é, que não estavam realizando compras parceladas quando do lançamento da campanha de marketing. Não foi realizada nenhuma segmentação adicional e as personalizações da carta resumiam-se ao nome do cliente e ao crédito que lhe era oferecido.

4.4. BASE DE DADOS

A Empresa X disponibilizou duas bases para serem utilizadas nesta tese. A primeira contém registros das transações efetuadas por todos os cliente que fazem parte do público-alvo da campanha. A segunda contém dados relativos ao resultado da campanha.

Cada registro da base de transações representa uma transação realizada por um cliente. A base de dados que foi disponibilizada possui 138.354 registros, que representam todas as transações realizadas entre Janeiro de 2000 e Outubro de 2002 pelos clientes que fazem parte do público-alvo da campanha. A base de transações possui o seguinte *layout*:

Coluna	Descrição
Identificador de Cliente	Código identificador do cliente na base de dados.
Data da Compra	Data em que a compra foi realizada pelo cliente.
Ticket da Compra	Valor pago pelo cliente.

Tabela XII: *Layout* da base de transações.

Na base de resultado da campanha estão representados a reação à campanha de todos os clientes que foram selecionados para fazer parte do público alvo e, portanto, receberam a comunicação da oferta (mala-direta da campanha). Em outras palavra, nesta base, além do identificador do cliente, estão presentes uma variável que indica se o cliente comprou ou não um

produto como resultado da campanha e o valor gasto na compra destes produtos.

A base de dados possui, naturalmente, 50.529 registros, sendo que 985 registros possuem o indicador “Comprou?” marcado com o número “1”, significando que o cliente realizou uma compra motivado pela campanha de marketing direto. Esta base de dados possui o seguinte *layout*:

Coluna	Descrição
Identificador de Cliente	Código identificador do cliente na base de dados.
Comprou?	Indica se o cliente realizou alguma compra motivado pela campanha de marketing direto.
Ticket da Campanha	Valor pago pelo cliente.

Tabela XIII: Layout da base de resultado da campanha.

4.5. IMPLEMENTAÇÃO DO RFV OTIMIZADO

O modelo de otimização para a análise RFV foi implementado em uma planilha do Microsoft Excel com auxílio do software de otimização, baseado em Algoritmo Genético, Evolver for Excel, da Palisade Corporation (www.palisade.com).

4.5.1. UTILIZANDO O MS EXCEL

Para a construção do modelo de otimização da análise RFV foi utilizada uma amostra típica com 3.000 clientes. Esta amostra extraída aleatoriamente foi armazenada na planilha Excel, juntamente com as informações de classificação RFV e resultado de campanha.

A planilha foi construída na versão 2000 do MS Excel e está dividida em cinco pastas: Parâmetros, Registros Transações, Registros no Período, Campanha e Resultado.

4.5.1.1. PASTA DE PARÂMETROS

Na primeira pasta estão registrados os parâmetros globais de operação do modelo de otimização para análise RFV, isto é:

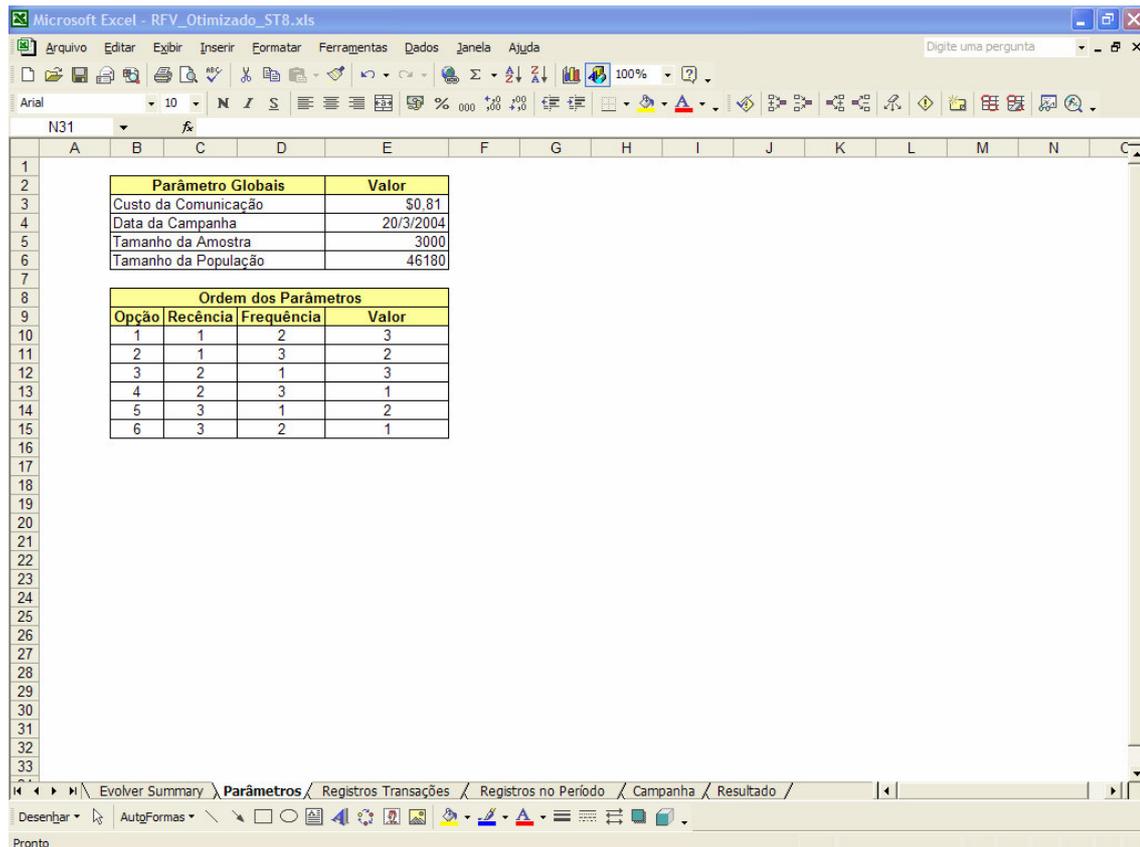
- Custo da comunicação – representa o custo unitário de produção e envio das peças de comunicação da campanha de marketing direto.

Este custo precisa ser retirado do ticket da transação realizada por um cliente para que seja calculado o resultado da campanha. Clientes que receberam a comunicação mas não responderam positivamente à oferta terão este custo como resultado individual da campanha.

- Data da Campanha – data de início da campanha, ou seja, data em que as comunicações foram enviadas para os clientes do público alvo.
- Tamanho da Amostra – quantidade de registros da amostra retirada da base de dados disponibilizada pela Empresa X para construção do modelo.
- Tamanho da População – quantidade de registros da base de dados disponibilizada pela Empresa X para construção do modelo.
- Ordem dos Parâmetros – tabela de configuração que determina a ordem dos parâmetros de recência, frequência e valor no RFV Otimizado.

A primeira coluna da tabela (“opção”) estará associada a uma das variáveis manipuladas pelo processo de otimização. As demais

colunas determinam a ordem de cada parâmetro. Por exemplo, se o Algoritmo Genético selecionar a opção 1, a ordem das dimensões será recência, freqüência e valor. Se o Algoritmo Genético selecionar a opção 2, a ordem das dimensões será recência, valor e freqüência.



The screenshot shows the 'Parâmetros' worksheet in Microsoft Excel. It contains two tables:

Parâmetro Globais		Valor
Custo da Comunicação		\$0.81
Data da Campanha		20/3/2004
Tamanho da Amostra		3000
Tamanho da População		46180

Ordem dos Parâmetros			
Opção	Recência	Freqüência	Valor
1	1	2	3
2	1	3	2
3	2	1	3
4	2	3	1
5	3	1	2
6	3	2	1

Figura XIV: Pasta “Parâmetros” na planilha do RFV Otimizado

A pasta de parâmetros está apresentada na figura XIV.

4.5.1.2. PASTA DE REGISTROS DE TRANSAÇÃO

A Pasta de Registros de Transação contém a amostra retirada da base com as transações de clientes disponibilizada pela Empresa X, sendo que cada registro representa uma transação realizada pelo cliente da Empresa X no período de 2000 a 2002. Os campos armazenados na Pasta de Registros de Transação são os seguintes:

- # caso – número seqüencial que identifica o registro na planilha.
- Identificador do Cliente – código identificador do cliente na base de dados da empresa.
- Data da Compra – data em que o cliente realizou a compra em questão na Empresa X.
- Ticket da Compra – valor pago pelo cliente na compra de mercadorias na Empresa X.

# caso	Ident. Cliente	Data da Compra	Ticket da Compra	Diferença Dias Recência
1				317
2				812
3				311
4				803
5				792
6				792
7				792
8				792
9				792
10				792
11				792
12				749
13	151782	13/12/2000	\$ 10,00	718
14	151782	13/12/2000	\$ 448,47	718
15	151782	13/12/2000	\$ 189,00	718
16	404277	21/12/2000	\$ 788,60	710
17	404277	21/12/2000	\$ 742,48	710
18	2003869	14/01/2001	\$ 227,76	686
19	2003869	14/01/2001	\$ 256,96	686
20	2043579	16/01/2001	\$ 466,36	684
21	2043579	16/01/2001	\$ 758,16	684
22	1931354	20/01/2001	\$ 470,49	680
23	2692616	10/02/2001	\$ 799,00	659
24	3113311	15/03/2001	\$ 25,79	626
25	3113311	15/03/2001	\$ 73,70	626
26	3113311	15/03/2001	\$ 159,34	626
27	3113311	15/03/2001	\$ 323,63	626
28	2962294	15/03/2001	\$ 66,126	626
29	3117901	17/03/2001	\$ 340,84	624
30	3118719	17/03/2001	\$ 347,58	624
31	2003869	17/03/2001	\$ 902,88	624
32	2003869	08/04/2001	\$ 588,80	602
33	6712358	14/05/2001	\$ 299,00	566
34	6676904	02/06/2001	\$ 389,00	547
35	6074442	04/06/2001	\$ 299,00	545
36	3813044	06/06/2001	\$ 188,79	543
37	3822539	06/06/2001	\$ 149,00	543
38	3822539	06/06/2001	\$ 149,00	543
39	3822539	06/06/2001	\$ 75,00	543
40	9086493	06/06/2001	\$ 413,36	543

Figura XV: Pasta “Registros no Período” na planilha do RFV Otimizado

4.5.1.3. PASTA DE REGISTROS NO PERÍODO

A Pasta de Registros no Período é semelhante à pasta de Registros de Transações, pois os registros da primeira que estão dentro do período configurado pelo processo de otimização são copiados para esta pasta. Por

isso as colunas “# caso”, “Identificação do Cliente”, “Data da Compra” e “Ticket da Compra”, previamente descritas, repetem-se nesta pasta.

A coluna “Diferença Dias Recência”, introduzida nesta pasta, é calculada como a diferença, em dias, da data da campanha para a data da última transação realizada pelo cliente. Se esta diferença for maior do que a configurada pelo processo de otimização, o registro não é copiado para esta pasta. A pasta de Registros no Período está apresenta na Figura XV.

4.5.1.4. PASTA DE CAMPANHA

Na Pasta de Campanha estão os registros das comunicações realizadas pela Empresa X na campanha de marketing direto utilizada na tese. Cada registro representa uma comunicação enviada para um dos clientes. Os campos armazenados na Pasta de Campanha são os seguintes:

- Dados do Caso (campanha) – são as informações da campanha realizada pela Empresa X:
 - # caso – número seqüencial que identifica o registro na planilha.
 - Identificador do Cliente – código identificador do cliente na base de dados da empresa.
 - Comprou? – indica se o cliente realizou alguma compra motivado pela campanha de marketing direto. Quando a compra foi realizada a variável recebe o valor “1” e quando a compra não foi realizada a variável recebe o valor “0”.
 - Ticket – valor da compra realizada pelo cliente em decorrência da campanha. Quando o cliente não respondeu positivamente à campanha, ou seja, não realizou a compra, esta coluna possui o valor zero.

- Resultado Real – cálculo do valor do ticket subtraído o custo de comunicação que está parametrizado na Pasta de Parâmetros.
- Dados RFV – são os dados necessários à classificação RFV do cliente.
 - Ticket Anterior – soma de todas as transações realizadas pelo cliente no período anterior à campanha que estejam dentro do período configurado pelo processo de otimização.
 - Data da Última Compra – data da transação mais recente realizada pelo cliente no período anterior à campanha de marketing direto.
 - Quantidade de Dias de Recência – quantidade de dias entre a data da campanha e a última transação realizada pelo cliente.
 - Quantidade de Transações – quantidade de transações realizadas pelo cliente no período anterior à campanha que estejam dentro do período configurado pelo processo de otimização.
- Classificação RFV – dados utilizados durante o processo de classificação RFV do cliente.
 - Índice na Primeira Dimensão – posição relativa da primeira dimensão do cliente no conjunto de todos os clientes do público-alvo. A primeira dimensão poderá receber os dados de recência, frequência ou valor de acordo com a opção configurada pelo algoritmo de otimização.
 - Primeira Dimensão – classificação do cliente na primeira dimensão de classificação.

- Índice na Segunda Dimensão – posição relativa da segunda dimensão do cliente no conjunto de todos os clientes do público-alvo. A segunda dimensão poderá receber os dados de recência, frequência ou valor de acordo com a opção sorteada pelo algoritmo de otimização.
- Segunda Dimensão – classificação do cliente na segunda dimensão de classificação.
- Índice na Terceira Dimensão – posição relativa da terceira dimensão do cliente no conjunto de todos os clientes do público-alvo. A terceira dimensão poderá receber os dados de recência, frequência ou valor de acordo com a opção sorteada pelo algoritmo de otimização.
- Terceira Dimensão – classificação do cliente na terceira dimensão de classificação.
- Grupo - classificação final do cliente nas três dimensões na ordem configurada pelo processo de otimização.

Além destas colunas, nesta pasta existe um *box* de escolhas onde estão os parâmetros ajustáveis do RFV Otimizado, a saber: Dimensões de Valor, Dimensões de Recência, Dimensões de Frequência, Ordem e Período. O processo de otimização realizará as alterações nestas variáveis e avaliará o resultado obtido na Pasta de Resultado.

4.5.1.5. PASTA DE RESULTADO

Na Pasta de Resultado é estimado o resultado da campanha de marketing direto com as configurações da classificação RFV obtidas a partir dos parâmetros inseridos pelo processo de otimização na Pasta de Campanha. Nesta pasta existem as seguintes informações:

- Classificação – classificação RFV que identifica um grupo de clientes.
- Resultado por Classificação – resultado da campanha no grupo de clientes identificado na coluna “Classificação”. Este resultado é calculado com base no resultado individual de cada cliente que teve o seu registro considerado por estar dentro do período de corte.
- Envia? – indicador que define se a comunicação deve ser ou não enviada para os clientes do grupo com base no resultado estimado.

Foi adotada a política de que um grupo de clientes não deveria receber a comunicação se o seu resultado estimado apontar um possível prejuízo da campanha no grupo em questão.

The screenshot shows a Microsoft Excel spreadsheet titled "RFV_Otimizado_7.xls". The spreadsheet is divided into two main sections: "Dados do Caso (campanha)" and "Dados RFV".

Dados do Caso (campanha)					Dados RFV					Clas	
# caso	Ident. Cliente	comprou?	ticket	Res. Real	Ticket Anterior	Dt. Ult. Compra	Qtz. Dias Rece.	Qtz. Transações	Índice Recência	Recência	Índice Freq
1	75145	0	\$ -	(\$2,00)	\$ 2.491,48	26/08/02	97	8	8%		1
2	76290	0	\$ -	(\$2,00)	\$ 210,00	11/05/01	569	1	82%		5
3	76954	0	\$ -	(\$2,00)	\$ 1.154,92	12/05/01	568	3	80%		4
4	82669	0	\$ -	(\$2,00)	\$ 3.177,41	13/08/02	110	16	15%		1
5	84285	0	\$ -	(\$2,00)	\$ 5.478,60	07/10/02	55	9	3%		1
6	84365	0	\$ -	(\$2,00)	\$ 1.261,51	02/09/02	90	5	6%		1
7	88199	0	\$ -	(\$2,00)	\$ 469,70	04/04/01	606	1	86%		5
8	98680	0	\$ -	(\$2,00)	\$ 1.511,30	02/02/02	302	8	48%		3
9	98692	0	\$ -	(\$2,00)	\$ 901,80	13/04/02	232	6	33%		2
10	99290	0	\$ -	(\$2,00)	\$ 3.061,56	11/12/01	355	12	52%		3
11	99609	0	\$ -	(\$2,00)	\$ 1.736,93	06/04/02	239	6	34%		2
12	99747	0	\$ -	(\$2,00)	\$ 1.398,31	29/10/02	33	3	0%		1
13	100047	0	\$ -	(\$2,00)	\$ 841,24	14/12/01	352	2	51%		3
14	102010	0	\$ -	(\$2,00)	\$ 1.096,44	25/06/01	524	2	76%		4
15	103631	0	\$ -	(\$2,00)	\$ 1.613,15	08/03/02	268	2	39%		2
16	111028	0	\$ -	(\$2,00)	\$ 3.073,86	28/11/01	368	7	56%		3
17	112909	0	\$ -	(\$2,00)	\$ 1.001,02	06/02/02	298	3	46%		3
18	118267	0	\$ -	(\$2,00)	\$ 349,48	18/03/02	258	2	37%		2
19	119952	0	\$ -	(\$2,00)	\$ 968,18	02/08/02	121	3	23%		2
20	120475	0	\$ -	(\$2,00)	\$ 3.760,53	20/08/02	103	13	10%		1
21	120886	0	\$ -	(\$2,00)	\$ 751,16	04/02/02	300	4	47%		3
22	122099	0	\$ -	(\$2,00)	\$ 922,56	21/07/01	498	3	74%		4
23	123608	0	\$ -	(\$2,00)	\$ 1.449,64	03/12/01	363	2	53%		3
24	123948	0	\$ -	(\$2,00)	\$ 569,00	26/04/02	219	2	31%		2
25	127229	0	\$ -	(\$2,00)	\$ 2.338,56	06/10/01	421	3	62%		4
26	127279	0	\$ -	(\$2,00)	\$ 2.153,70	09/10/01	418	4	61%		4
27	127819	0	\$ -	(\$2,00)	\$ 1.382,60	03/08/02	120	9	22%		2
28	127976	0	\$ -	(\$2,00)	\$ 3.220,12	05/10/01	422	8	63%		4
29	131279	0	\$ -	(\$2,00)	\$ 838,00	08/07/02	146	5	25%		2
30	131616	0	\$ -	(\$2,00)	\$ 1.005,80	28/11/01	368	3	56%		3
31	139757	0	\$ -	(\$2,00)	\$ 976,23	14/11/00	747	1	88%		5
32	140135	1	\$ 124,82	\$ 122,82	\$ 299,00	09/11/00	752	1	94%		5

Figura XVI: Pasta “Campanha” na planilha do RFV Otimizado

Classificação RFV	Índice Freqüência	Índice de Valor	Valor	Valor	Classificação Final	Freqüência	Recência	Grupo
76%	4	75%	4	4	4	1	1	441
0%	1	0%	1	1	1	5	1	115
27%	2	39%	2	2	2	4	2	224
100%	5	85%	5	5	5	1	1	551
83%	5	99%	5	5	5	1	1	551
53%	3	45%	3	3	3	1	1	331
0%	1	4%	1	1	1	5	1	115
76%	4	58%	3	3	4	3	3	343
61%	4	21%	2	2	4	2	2	242
93%	5	83%	5	5	5	3	3	553
61%	4	64%	4	4	4	2	2	442
27%	2	51%	3	3	2	1	1	321
10%	1	10%	1	1	1	3	3	113
10%	1	37%	2	2	1	4	4	214
10%	1	61%	4	4	1	2	2	412
68%	4	84%	5	5	4	3	3	543
27%	2	29%	2	2	2	3	3	223
10%	1	3%	1	1	1	2	2	112
27%	2	24%	2	2	2	2	2	222
96%	5	90%	5	5	5	1	1	551
39%	2	16%	1	1	2	3	3	123
27%	2	23%	2	2	2	4	4	224
10%	1	55%	3	3	1	3	3	313
10%	1	6%	1	1	1	2	2	112
27%	2	72%	4	4	2	4	4	424
39%	2	70%	4	4	2	4	4	424
83%	5	49%	3	3	5	2	2	352
76%	4	86%	5	5	4	4	4	544
53%	3	18%	1	1	3	2	2	132
27%	2	30%	2	2	2	3	3	223
0%	1	25%	2	2	1	5	5	215
0%	1	2%	1	1	1	5	5	115

Escolhas	
Dim. Valor	5
Dim. Recência	5
Dim. Freqüência	5
Ordem	6
Período	718

Figura XVII: Pasta “Campanha” na planilha do RFV Otimizado

- Resultado se Enviou – caso, pela política definida, se decida que o grupo de clientes deve ser contatado, ou seja, devem ser enviadas as comunicações com a oferta da campanha, nesta coluna fica representado o resultado estimado do grupo na campanha. Sumarizando-se estas colunas obtém-se a estimativa de resultado da campanha com a utilização do RFV Otimizado.

4.5.2. UTILIZANDO O EVOLVER FOR EXCEL

O Evolver for Excel, lançado em 1989, é o primeiro software disponível comercialmente para o desenvolvimento de aplicações com Algoritmos Genéticos. Trata-se de uma ferramenta desenvolvida pela empresa *Palisade Corporation* (PALISADE, 2001).

4.5.2.1. PORQUE UTILIZAR EVOLVER

A seleção do software Evolver para implementação da otimização baseada em Algoritmos Genéticos baseou-se nas seguintes vantagens oferecidas pela ferramenta:

- A versão industrial do Evolver admite um número ilimitado de variáveis;
- O Evolver é capaz de lidar com problemas lineares e não lineares, incluindo-se aqui os problemas combinatórios e de permutação de valores;
- As funcionalidades providas pelo MS Excel facilitam, em muito, a descrição de problemas cuja solução se deseja otimizar;
- A taxa de mutação empregada em Algoritmos Genéticos é ajustada dinamicamente pelo Evolver, evitando que o mecanismo de

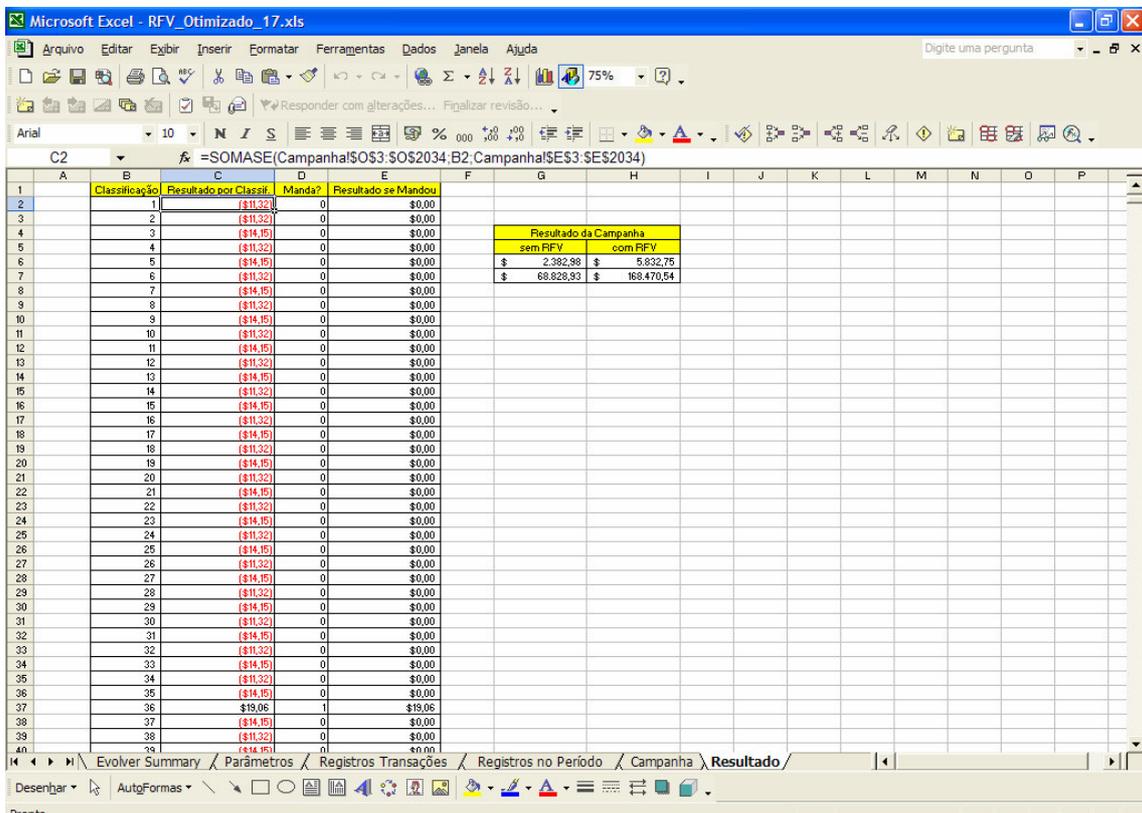


Figura XVIII: Pasta "Resultado" na planilha do RFV Otimizado

otimização fique preso em pontos ótimos locais;

- O Evolver oferece diversos relatórios da evolução do mecanismo de otimização cujas informações são atualizadas dinamicamente, como por exemplo, o gráfico de progresso da otimização com o resultado da melhor solução encontrada e a solução média, a tabela de cores que representa a diversidade dos organismos utilizados no processo de otimização e a tabela de valores dos genes, com os valores sorteados e o resultado obtido em cada gene da população;
- O Evolver oferece seis algoritmos diferentes para facilitar a modelagem dos mais diversos tipos de problema, a saber:
 - *Recipe* – indicado para ser utilizado quando o conjunto de variáveis cujos os valores podem ser alterados pelo mecanismo de otimização podem variar livremente, da mesma forma que os ingredientes de uma receita culinária podem ser combinados com o objetivo de se determinar a mistura mais adequada à receita;
 - *Grouping* – indicado para ser utilizado quando elementos no contexto do problema têm que ser acondicionados em grupos. Por exemplo, muitas vezes trabalhadores precisam ser organizados em grupos para a execução de uma determinada tarefa, como na linha de montagem da indústria automobilística;
 - *Order* – indicado para ser utilizado quando se precisa determinar a melhor ordem possível para a execução de tarefas independentes necessárias à solução de um determinado problema. Por exemplo, a organização ótima das mercadorias nas gôndolas de um supermercado;
 - *Budget* – indicado para ser utilizado quando o conjunto de variáveis cujos os valores podem ser alterados pelo

mecanismo de otimização podem variar livremente restritos a uma condição comum. Por exemplo, a distribuição do orçamento anual de uma empresa entre os diversos departamentos. Cada departamento pode receber qualquer parcela do orçamento anual desde que a soma dos orçamentos de todos os departamentos não ultrapasse o total determinado pela empresa;

- *Project* – indicado para ser utilizado quando se precisa determinar a melhor ordem possível para a execução de tarefas interdependentes necessárias à solução de um determinado problema. Por exemplo, a organização ótima das tarefas no contexto de um projeto de desenvolvimento de sistemas de informação;
- *Schedule* – indicado para ser utilizado quando elementos no contexto do problema têm que ser acondicionados em blocos de tempo respeitando-se determinadas restrições. Por exemplo, a organização dos turnos de trabalho em uma companhia aérea;

Tudo isso torna o Evolver um software extremamente atrativo para ser utilizado em problemas de otimização. Todas estas funcionalidades podem ser adquiridas por um preço inferior a US\$ 1,000. Isso significa que este software oferece todas estas funcionalidades importantes a um preço vantajoso em relação a outras ferramentas do mercado, tais como:

- Premium Solver Platform (www.solver.com), oferecida por US\$ 2,545.00;
- FlexTool for MATLAB environment , que pode ser comprada por US\$ 999.00. Esta ferramenta, porém, precisa do software MATLAB para operar, o que adiciona um investimento de US\$ 1,000.00 para

comprar o software somados aos custos de treinamento e ambientação do MATLAB;

4.5.2.2. COMO O EVOLVER FUNCIONA

O Evolver é uma ferramenta de otimização baseada em Algoritmos Genéticos que funciona em conjunto com o Excel. O Excel provê as facilidades para a utilização de funções e fórmulas necessárias à descrição do problema cuja solução se deseja otimizar, enquanto o Evolver provê o mecanismo de execução de um Algoritmo Genético de propriedade da *Palisade Corporation* (www.palisade.com) e a interface de configuração deste mecanismo.

O processo de otimização baseada em Algoritmo Genético utilizando o Evolver começa com a correta representação do problema que se deseja otimizar em uma planilha Excel. A planilha contém, na verdade, um modelo do problema em questão. O modelo deve conter necessariamente um conjunto de variáveis de entrada (que são alteradas pelo mecanismo de execução de Algoritmo Genético), chamadas “células ajustáveis”, e produzir um resultado, que será avaliado pela função de *fitness*.

Quando o Evolver é iniciado uma barra de ferramentas é apresentada com botões que permitem a configuração do mecanismo de execução do Algoritmo Genético e o início, pausa ou paralisação do processo de otimização. Ver Figura XIX.

No módulo de configuração do Evolver é possível se selecionar a “célula-alvo” (*For the cell*). Esta é a célula que abriga o resultado do modelo que o mecanismo de Algoritmo Genético está buscando otimizar. É necessário definir se o objetivo desejado é maximizar (*Find the – Maximum*), minimizar (*Find the – Minimum*) ou aproximar o resultado do modelo de um dado valor (*Find the – Closest Value to*). Ver Figura XX.

Neste mesmo módulo definem-se as células que sofrerão os ajustes (*By Adjusting the Cells*) realizados pelo mecanismo de Algoritmo Genético, o Evolver chama estas células de “células ajustáveis” (*adjustable cells*), e as regras a que estes ajustes devem respeitar (*Subject to the Constraints*). Ao se adicionar uma célula ajustável, o Evolver apresenta o módulo de detalhe das células ajustáveis onde o usuário deve definir o algoritmo de solução dentre os disponíveis (*Combo box: Use the “RECIPE” Solving Method*), detalhar as regras às quais as células ajustáveis estão sujeitas (*Adjust The Cells*), a taxa de *crossover* (*Crossover rate*) e mutação (*Mutation rate*), além de uma descrição significativa que auxilie na identificação das células ajustáveis (*Description*). Ver Figura XXI.



Figura XIX: Controle das operações do Evolver.

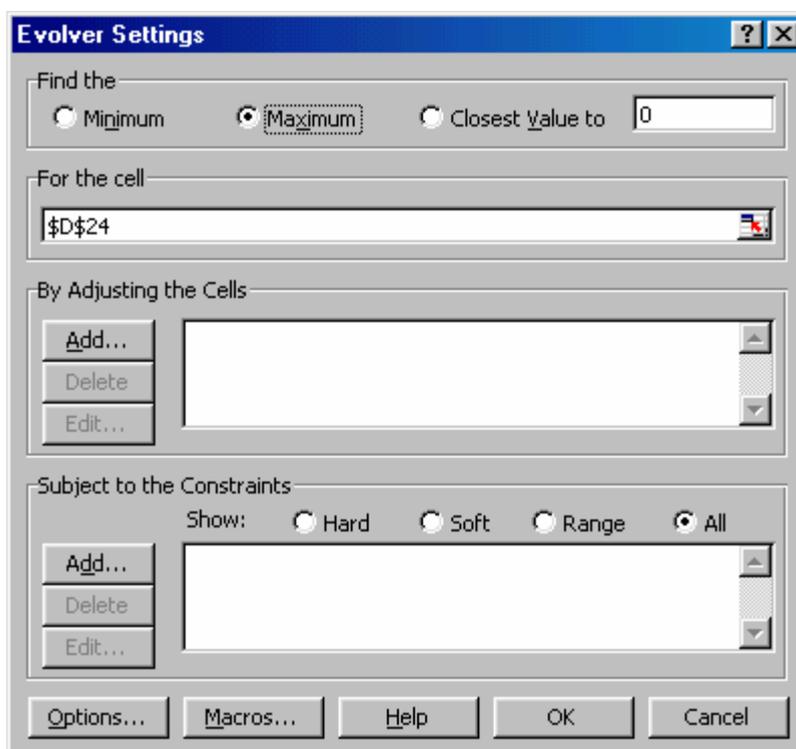


Figura XX: Módulo de configuração do Evolver

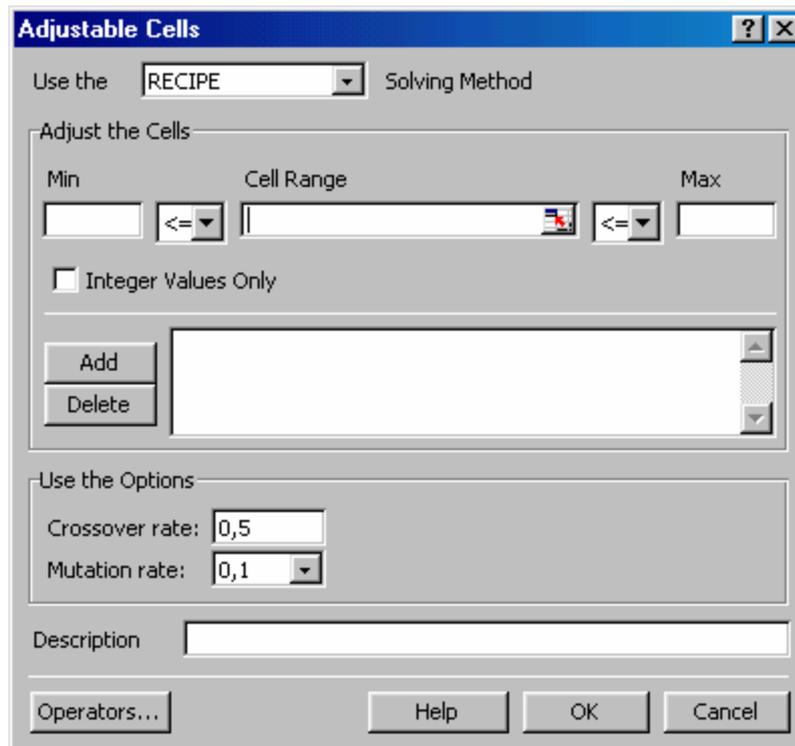


Figura XXI: Opções das células ajustáveis

O Evolver possui, ainda, um módulo de definição de opções adicionais que são: o tamanho inicial da população (*Population Size*), a opção de pausar a otimização em caso de erro (*Pause on Error*), de apresentar o gráfico de progresso (*Graph Progress*), atualizar a planilha apresentada a cada iteração (*Update the Display – Every Calculation*), quando o resultado for melhor que o anterior (*Update the Display – With Only The Best Result*) ou apenas no final do processo (*Update the Display – Never*), se os valores iniciais são aleatórios (*Random Number Seed – Generated Randomly*) ou fixos (*Random Number Seed – Fixed*) e as condições de parada (*Stopping Conditions*). Ver Figura XXII.

Existem várias opções de condição de parada que podem ser utilizadas em conjunto, a saber:

- Número fixo de tentativas (*Trials*) – o processo de otimização termina quando um número determinado de iterações é atingido;

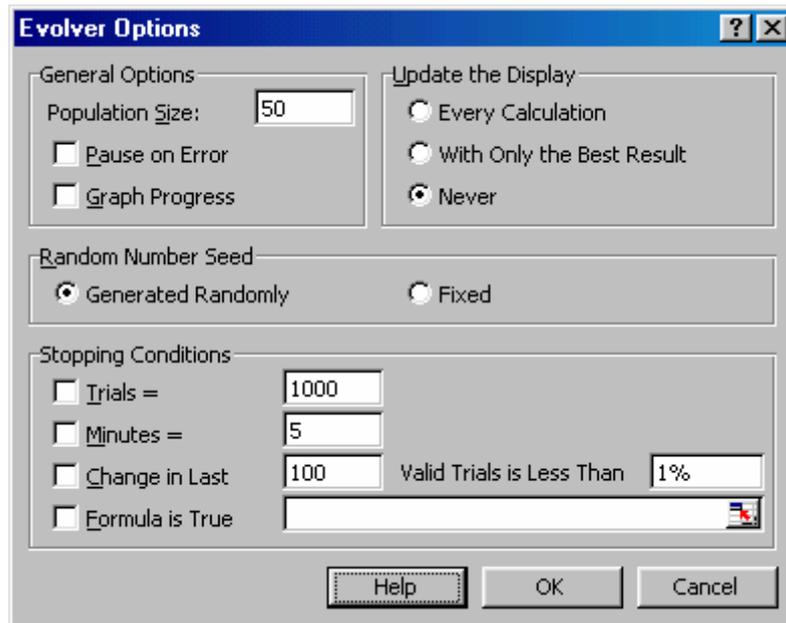


Figura XXI: Módulo de Opções do Evolver.

- Tempo de processamento (*Minutes*) – o processo permanece rodando por um período máximo determinado;
- Percentual de mudança nas últimas iterações (*Change in Last n Valid Trials is Less Than m%*) – o processo permanece rodando enquanto o percentual de mudança na últimas n iterações é maior que $m\%$, sendo “ n ” e “ m ” informados pelo usuário; e
- Fórmula é verdadeira (*Formula is True*) – uma fórmula implementada em uma célula da planilha retorna um valor verdadeiro.

Durante a execução do processo de otimização é possível acompanhar sua evolução no Evolver através do módulo “*Evolver Watcher*” e através da própria planilha onde o problema está representado. Na janela “*Evolver Watcher*” são apresentados os relatórios de evolução que são atualizados dinamicamente. Ver figura XXIII.

4.5.3. CRIANDO UM MODELO DE OTIMIZAÇÃO RFV NO EVOLVER

As opções de configuração utilizadas para criar o modelo de RFV Otimizado, apresentado no item 4.5.1., no Evolver for Excel são descritas a seguir. Estas configurações envolvem, na ordem:

- Configuração das células ajustáveis;
- Configuração das opções do Evolver;

4.5.3.1. CONFIGURANDO AS CÉLULAS AJUSTÁVEIS

As configurações das células ajustáveis é apresentada na Figura XXIV. As descrições dos parâmetros utilizados para configurar as células ajustáveis são apresentadas na tabela XIV.

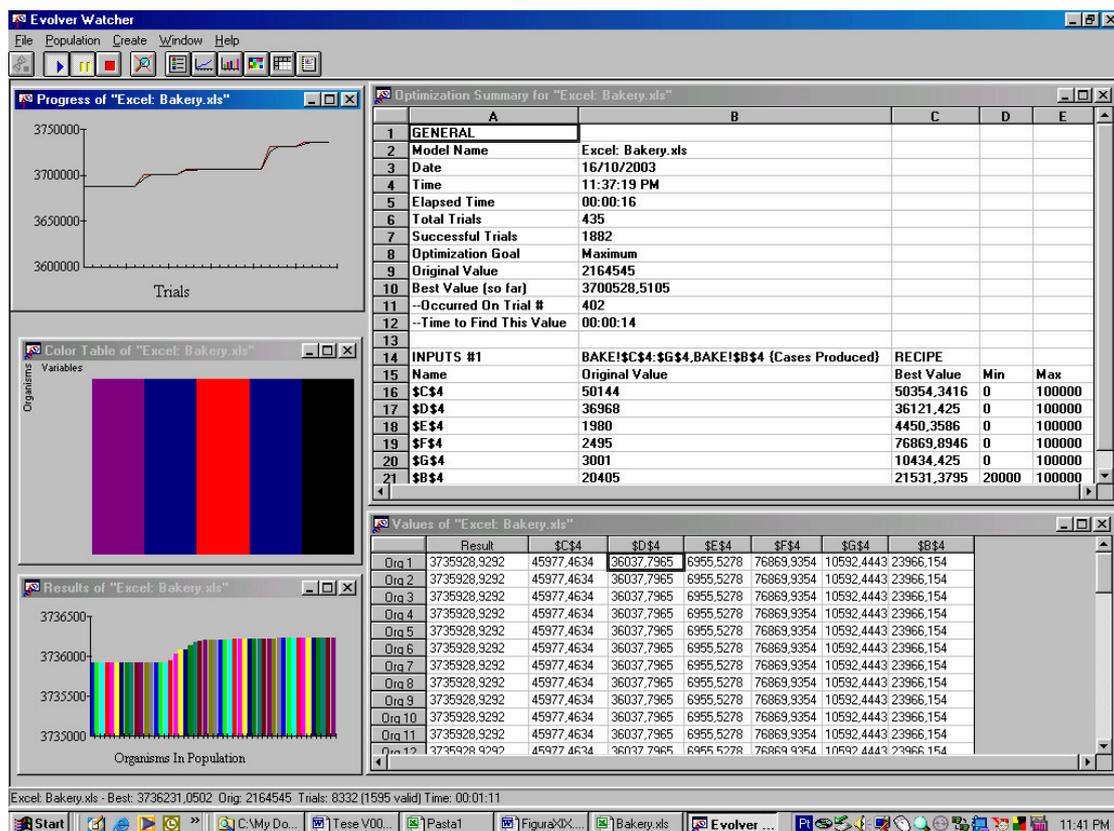


Figura XXIII: Relatórios do Evolver.

Configuração	Valor	Descrição
<i>Find the</i>	<i>Maximum</i>	Deseja-se encontrar a configuração do RFV Otimizado que maximiza o resultado estimado da campanha de marketing direto.
<i>For the cell</i>	\$W\$67	Esta é a função de aptidão deste modelo, representada pela célula onde é calculado o resultado estimado do modelo com base no resultado da amostra a partir das configurações de RFV realizadas pelo mecanismo de otimização.
<i>By adjusting the Cells</i>	\$V\$2 {Dimensão de Valor} \$V\$3 {Dimensão de Recência} \$V\$4 {Dimensão de Frequência} \$V\$5 {Ordem} \$V\$6 {Período}	Estas são as células que podem ser ajustadas para que o modelo calcule o resultado estimado da campanha de marketing. Representam as decisões de classificação RFV.
<i>Subject to the Constraints</i>	R: 1<=\$V\$2<=1000[INT] R: 1<=\$V\$3<=1000[INT] R: 1<=\$V\$4<=1000[INT] R: 1<=\$V\$5<=6[INT] R: 1<=\$V\$6<=890[INT]	São as regras a que estão submetidos os ajustes que as células de configuração poderão receber. Por exemplo, a célula \$V\$2 poderá conter valores inteiros de 1 a 1000, enquanto a célula \$V\$5 valores inteiros de 1 a 6.

Tabela XIV: Descrição dos parâmetros utilizados nas células ajustáveis

Observe que o método “RECIPE” foi utilizado em todas as células ajustáveis. Além deste ser o método mais popularmente utilizado (PALISADE, 2001), ele encaixa-se perfeitamente na necessidade do modelo, onde as variáveis podem variar livremente sem interdependência entre elas. Observe também que optou-se pelos valores padrão do Evolver for Excel para a taxa de *crossover* (0,5) e para a taxa de mutação (0,1).

4.5.3.2. CONFIGURANDO AS OPÇÕES DO EVOLVER

Após a configuração das células ajustáveis e da célula que se deseja otimizar (célula com o resultado do modelo) é necessário configurar os parâmetros de funcionamento do software de algoritmo genético. As configurações do mecanismo de otimização do Evolver for Excel são apresentadas na Figura XXIV. As descrições dos parâmetros utilizados para configurar o mecanismo são apresentadas na tabela XV.

Configuração	Valor	Descrição
<i>Population Size</i>	50	Tamanho inicial da população gerado aleatoriamente. Em geral uma população de 50 organismos é suficiente para problemas medianamente complexos. Neste caso, o tamanho da população mostrou-se mais adequado.
<i>Pause on Error</i>	SIM	No caso das fórmulas da planilha gerarem algum erro durante o processo, este deverá ser paralisado.
<i>Graph Progress</i>	SIM	Indica que desejamos acompanhar o gráfico de progresso durante o processo de otimização.
<i>Update the Display</i>	<i>With only The Best Results</i>	Determina que a planilha seja atualizada apenas quando o mecanismo de otimização selecionar uma configuração com resultados melhores do que os que estão apresentados na planilha. Esta opção melhora a performance do processo.
<i>Random Number Seed</i>	<i>Generated Randomly</i>	Neste caso optou-se pela geração aleatória dos organismos iniciais sempre que o processo for iniciado.
<i>Stopping Condition</i>	<i>Change in Last 200 Valid Trials is Less Than 1%</i>	Esta condição de parada determina que o processo de otimização permaneça rodando até que a melhora de resultado nas últimas 200 tentativas válidas não seja maior que 1%.

Tabela XV: Descrição das opções utilizadas no modelo.

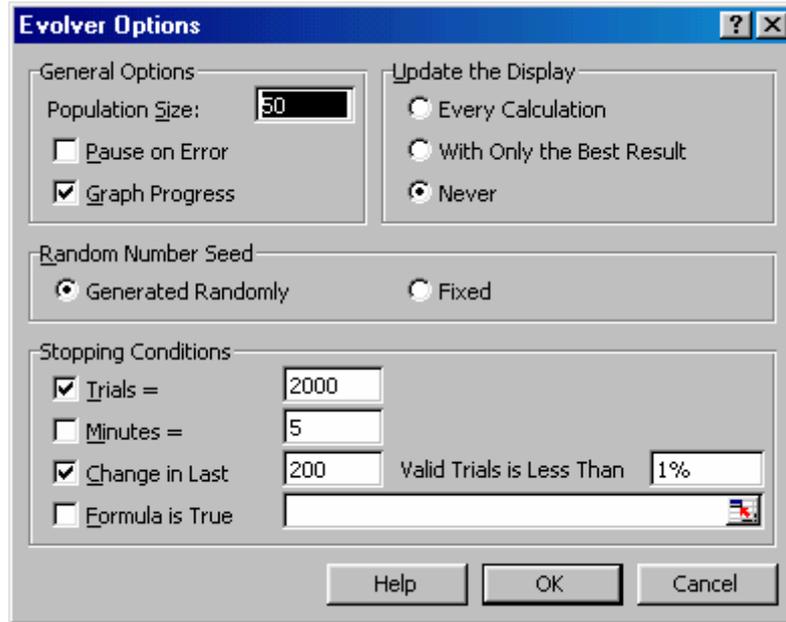


Figura XXV: Opções do modelo no Evolver.

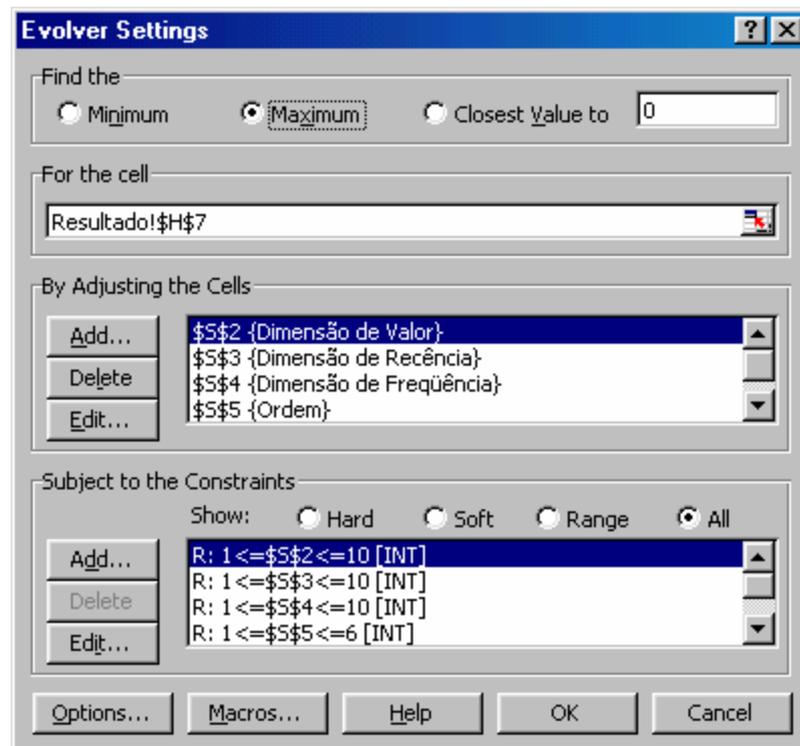


Figura XXIV: Configuração do modelo no Evolver.

Capítulo 5. ANÁLISE DOS RESULTADOS

5.1. INTRODUÇÃO

Neste capítulo são analisados os resultados da campanha de marketing direto executada pela Empresa X. Inicialmente, os resultados são analisados tendo em vista que a campanha foi realizada sem nenhum tipo de segmentação adicional. Posteriormente é feita a análise de qual seria o resultado, caso a Empresa X tivesse optado por utilizar a Análise RFV tradicional e o RFV Otimizado. Por último os resultados são comparados e analisados em conjunto.

5.2. RESULTADO REAL DA CAMPANHA

Para a realização da campanha de marketing direto objeto deste estudo a Empresa X não utilizou qualquer tipo de segmentação ou teste de mercado para estimar o resultado da campanha e enviar as comunicações apenas para os segmentos de clientes que, estima-se, reagiriam positivamente à oferta. Ao invés disso, a empresa optou por enviar a comunicação para todos os clientes do público-alvo.

Ainda sim, a campanha gerou uma receita bruta de R\$ 206.309,52 contra o investimento de R\$ 142.491,78 em custos de produção e envio das comunicações (custo unitário por comunicação de R\$ 2,83), produzindo um lucro, antes dos impostos, de R\$ 63.717,74, ou seja, 44,72% sobre o investimento.

É importante notar que este resultado foi obtido através de vendas efetuadas para apenas 985 cliente, ou seja, 1,95% dos clientes contatados. Este resultado está resumido na tabela XVI.

Dados da Campanha	Empresa X
Público-alvo (clientes)	50.529
Clientes do público-alvo que realização compras	985 (1,95%)
Custo de comunicação por cliente	R\$ 2,83
Custo total de comunicação	R\$ 142.491,78
Receita	R\$ 206.309,52
Lucro (Receita – Custo total de comunicação)	R\$ 63.817,74

Tabela XVI: Estatísticas da campanha de marketing direto da Empresa X.

O ticket médio da campanha, de R\$ 209,45, foi 16% abaixo do ticket médio da empresa. Outro ponto a ser observado é que 80% do faturamento total da campanha de marketing direto foi conquistado com as compras de 46,6% dos clientes que responderam positivamente à oferta. Estes números indicam que uma segmentação do público-alvo, com vista a eliminar segmentos que darão prejuízo à empresa, deve aumentar significativamente o resultado da campanha.

5.3. RESULTADO DA CAMPANHA COM ANÁLISE RFV

Neste item, é analisado o resultado da campanha de marketing direto realizada pela Empresa X, caso a Análise RFV tivesse sido aplicada da forma tradicional. As decisões de configuração da análise foram tomadas de forma arbitrária, utilizando-se o proposto por Hughes (HUGHES, 2000) e apresentadas no item 2.4 desta tese. Seguem os parâmetros configurados:

Parâmetro	Configuração
Número de dimensões de recência	5
Número de dimensões de frequência	5
Número de dimensões de valor	5
Ordem das Dimensões	RFV
Período de Análise	2 anos

Tabela XVII: Parâmetros do RFV.

Com esta configuração, os clientes foram classificados em 125 grupos e o resultado de cada grupo calculado, sendo que aqueles que tiveram um resultado financeiro negativo foram excluídos do público-alvo. A decisão de não enviar comunicações para os clientes destes grupos teria gerado uma economia total de R\$ 52.188,18 para a Empresa X.

Utilizando a Análise RFV, o faturamento bruto da campanha passou a ser de R\$ 181.167,15 enquanto o investimento foi de R\$ 90.303,60. Com isso o resultado final da campanha gerou um lucro antes dos impostos de R\$ 90.863,55. Neste caso, apenas 32.030 clientes receberam a comunicação da campanha de marketing direto.

5.4. RESULTADO DA CAMPANHA COM RFV OTIMIZADO

Neste item é aplicado aos dados disponibilizados pela Empresa X o modelo de otimização de resultado de campanha que busca a melhor configuração para a análise RFV. As configurações utilizadas na planilha Excel e no mecanismo de otimização estão apresentadas no Capítulo 4.

O mecanismo de otimização foi processado durante um período de aproximadamente doze horas e encontrou como configuração mais otimizada os seguintes parâmetros:

Parâmetro	Configuração
Número de dimensões de recência	8
Número de dimensões de freqüência	5
Número de dimensões de valor	17
Ordem das Dimensões	FVR
Período de Análise	830 dias

Tabela XVIII: Parâmetros do RFV Otimizado.

Com esta configuração de RFV, otimizada pelo modelo, os clientes foram separados em 680 grupos. Com esta classificação, a campanha de marketing direto geraria um resultado de R\$ 121.275,45 de lucro antes dos impostos.

É importante observar que o mecanismo de classificação configurou, para a campanha em questão, a dimensão de freqüência como sendo a mais importante na classificação dos clientes. Isso contraria o senso comum de que a recência deveria ser a primeira classificação por ser sempre a que possui maior poder de predição. Além disso, a separação dos clientes em oito grupos de recência e cinco grupos de freqüência e dezessete grupos de valor, para esta campanha, mostrou-se mais lucrativa.

5.5. COMPARAÇÕES

Com base nos resultados da campanha nas três modalidades apresentadas (sem utilizar a Análise RFV, com a Análise RFV tradicional e com o RFV Otimizado), é possível afirmar que a utilização do RFV Otimizado produziria um resultado maior para a Empresa X. Ver tabela XIX.

Quando comparados os resultados reais da campanha sem utilização de RFV, com os resultados após a aplicação da Análise RFV, observa-se um

crescimento de mais de R\$ 27.000,00 no lucro antes dos impostos. Este crescimento representa um aumento de 42,4% do lucro antes dos impostos.

Dados da Campanha		Sem RFV	RFV Tradicional	RFV Otimizado
Público-alvo (clientes)	Teste	-	3.000	3.000
	Rollout	50.529	29.030	21.224
	Total	50.529	32.030	27.224
Clientes que realizaram compras	Teste	-	65 (2,2%)	65 (2,2%)
	Rollout	985 (1,9%)	716 (2,5%)	703 (3,3%)
	Total	985 (1,9%)	781 (2,4%)	768 (2,8%)
Custos por comunicação	Teste	-	R\$ 2,83	R\$ 2,83
	Rollout	R\$ 2,83	R\$ 2,83	R\$ 2,83
Custos Totais de Comunicação	Teste	-	R\$ 8.460,00	R\$ 8.460,00
	Rollout	R\$ 142.491,78	R\$ 81.864,60	R\$ 59.851,68
	Total	R\$ 142.491,78	R\$ 90.303,60	R\$ 68.311,68
Receita	Teste	-	R\$ 12.811,80	R\$ 12.811,80
	Rollout	R\$ 206.309,52	R\$ 168.355,35	R\$ 176.775,33
	Total	R\$ 206.309,52	R\$ 181.167,15	R\$ 189.587,13
Lucro (Receita – Custo Total de Comunicação)		R\$ 63.817,74	R\$ 90.863,55	R\$ 121.275,45

Tabela XIX: Comparativo dos resultados da campanha de marketing direto.

Com a aplicação do RFV Otimizado o aumento do lucro é ainda maior, já que o resultado obtido com a configuração otimizada pelo modelo é de R\$ 121.275,45. Este valor representa um aumento de mais R\$ 30.000,00 no lucro

da campanha antes dos impostos em relação ao resultado da análise RFV. Este é um aumento de 33,5%.

Quando os resultados são comparados entre o RFV Otimizado e o resultado real da campanha, o incremento observado é ainda maior. A diferença monetária entre os dois resultados é de R\$ 54.457,71. Este valor representa um aumento de 90% no lucro da campanha. Os resultados podem ser comparados na figuras XV.

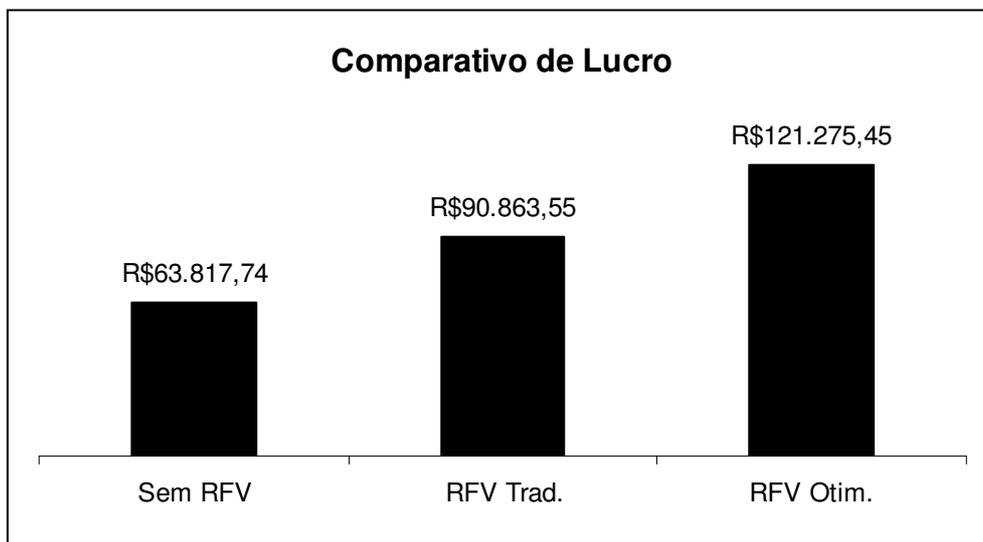


Figura XVI: Comparativo de lucro da campanha de marketing direto.

Capítulo 6. CONCLUSÕES

6.1. DISCUSSÃO

No início deste trabalho de tese nos propusemos a responder um conjunto de perguntas chaves. A respostas a estas perguntas são apresentadas abaixo:

P: Como o modelo que apresentamos maximiza os resultados de uma campanha de marketing direto?

R: Existem duas formas de se obter a maximização do retorno de uma campanha de marketing direto: redução de custos ou aumento da receita. Em ambos os casos a maximização do retorno é obtida através de uma série de decisões tomadas pelo gerente de marketing.

Para aumentar a receita de uma campanha de marketing o gerente responsável deve fazer o ajuste do *mix* de marketing, isto é, ajustar o preço, o produto e suas características, a estratégia de promoção e os elementos de distribuição às necessidades do público-alvo. Por exemplo, uma oferta direcionada à classe mais abastada pode se beneficiar da sua divulgação junto aos assinantes de revistas de grande circulação entre a parcela da população de maior poder aquisitivo, tais como Exame, Carta Capital, Forbes Brasil etc. Já o resultado de uma campanha direcionada à concessão de crédito pessoal de pequeno volume provavelmente se beneficiaria de um prazo de pagamento estendido, carência para início dos pagamentos, parcelas mensais reduzidas e divulgação junto aos clientes de baixo poder aquisitivo de uma cadeia de venda de eletrodomésticos.

Para reduzir os custos de uma campanha de marketing direto o gerente responsável deve tentar obter o menor custo possível na aquisição dos produtos que irão fazer parte da campanha, assim como dos produtos e

serviços que serão utilizados na implementação das estratégias de promoção e distribuição.

Por exemplo, em uma campanha de venda de CDs de musica junto aos assinantes de uma revista de grande circulação, o gerente de marketing pode optar por enviar as peças de comunicação da campanha junto com o exemplar da revista que o cliente recebe regularmente. O custo de distribuição dos CDs selecionados pelos clientes também pode ser reduzido se forem enviados junto com o próximo exemplar da revista.

Nesta tese apresentamos um modelo voltado para a diminuição de custos de campanhas de marketing direto. O modelo em questão pretende diminuir os custos indicando que segmentos de clientes provavelmente responderão de forma favorável ou desfavorável à oferta que estiver sendo realizada, ou seja, cujo o custo de comunicação, aquisição de produtos e distribuição são respectivamente menores ou maiores do que a receita a ser obtida.

Ao evitar contatar os segmentos de clientes que provavelmente não responderão favoravelmente, o gerente de marketing economiza não somente os custos de comunicação mas também os custos de compra de produtos, de estocagem e de distribuição.

P: Por que esforços devem ser empreendidos para otimizar a Análise RFV?

R: A Análise RFV é uma das técnicas mais populares de segmentação de clientes que fazem parte do conjunto de técnicas mais freqüentemente utilizadas pelos profissionais de marketing direto.

Ao oferecer um modelo de RFV otimizado esta tese está beneficiando uma parte expressiva de profissionais de marketing que atuam no mercado brasileiro e possibilitando que suas empresa tornem-se mais competitivas.

P: Por que a Análise RFV deve ser otimizada?

R: Para que a Análise RFV possa ser implementada, o gerente de marketing precisa determinar uma série de parâmetros utilizados no processo de análise, tais como período de tempo a ser considerado, número de segmentos de cada dimensão e a ordem em que as dimensões de recência, freqüência e valor devem ser aplicadas.

A determinação destes parâmetros é geralmente realizada de forma arbitrária com base em intuição e experiências pregressas, sendo que nem sempre esta escolha maximiza os resultados de uma campanha de marketing direto.

P: Como esta otimização pode ser realizada?

R: Obtendo a configuração ótima dos parâmetros que podem ser selecionados pelo gerente de marketing, isto é, o período de tempo, o número de segmentos em cada dimensão e a ordem em que as dimensões de recência, freqüência e valor devem ser aplicadas.

P: Como a Análise RFV Otimizada pode ser utilizada para tornar uma campanha de marketing direto mais lucrativa?

R: Para tornar uma campanha de marketing direto mais lucrativa a Análise RFV tradicional segmenta uma população de consumidores de acordo com a recência, freqüência e valor das compras efetuadas em um dado período de tempo. Desta população é retirada uma amostra que é submetida a uma oferta realizada por uma determinada organização.

A reação do consumidor à oferta efetuada pela organização permite a eliminação dos segmentos da população que, estima-se, não irão ultrapassar o ponto de *breakeven* da campanha de marketing.

A Análise RFV Otimizada identifica os segmentos que não irão ultrapassar o ponto de *breakeven* com uma precisão muito maior do que a Análise RFV tradicional, visto que a ordem em que as dimensões de recência, freqüência e valor são empregadas no processo de

segmentação (RFV, FRV, VRF, etc.), o número de segmentos em que cada dimensão será dividida e o período de tempo levado em consideração, são calculadas de forma a eliminar o maior número possível de segmentos que não irão ultrapassar o ponto de *breakeven* da campanha de marketing.

Observe que, no caso geral, em cada segmento existem tanto pessoas que vão reagir favoravelmente à oferta, efetuando a compra de produtos e serviços, quanto pessoas que irão reagir negativamente, declinando da oferta como um todo. A Análise RFV Otimizada separa, de forma ótima, um grupo do outro, permitindo a eliminação de um maior número de segmentos que não ultrapassarão o ponto de *breakeven*.

P: Como a Análise RFV Otimizada pode ser utilizada para determinar, com maior precisão, o perfil dos clientes que deverão fazer parte da campanha?

R: O perfil do cliente que irá reagir favoravelmente a uma campanha de marketing direto é determinado pela Análise RFV tradicional através da segmentação de uma população de consumidores de acordo com a recência, frequência e valor das compras efetuadas em um dado período de tempo.

Ao efetuar a separação ótima dos clientes que irão reagir favoravelmente a uma campanha dos demais, a Análise RFV Otimizada permite uma descrição mais precisa do perfil de recência, frequência e valor destes clientes, que pode ser enriquecido com outras informações que estejam disponíveis nos bancos de dados de consumidores, tais como, sexo, idade, classe social, grau de instrução, estado civil, etc.

P: Qual o impacto para as estratégias corporativas e de marketing de um modelo que maximiza os resultados das campanhas de marketing direto?

R: Campanhas mais lucrativas favorecem a utilização das técnicas de marketing direto em relação a outras alternativas que estejam à disposição do gerente de marketing. Como o marketing direto exige um conhecimento mais refinado das necessidades e características dos clientes de uma organização, o modelo que apresentamos acaba por favorecer a intensificação do relacionamento da organização com seus clientes, criando oportunidades para o desenvolvimento de produtos e serviços que estejam mais adequados às suas necessidades.

P: Como um modelo de otimização de resultado de campanha de marketing direto beneficia os clientes de uma organização e a sociedade em geral?

R: Ao tornar as empresas em operação no país mais competitivas através da redução dos custos de suas campanhas de marketing direto o modelo apresentado nesta tese favorece a atuação das forças de mercado no sentido de obter uma redução no preço final de venda de produtos e serviços em geral.

Ao evitar que peças de comunicação sejam enviadas para segmentos de clientes que iriam reagir desfavoravelmente a uma campanha de marketing direto o modelo que apresentamos melhora a qualidade da comunicação da empresa com seus clientes, aumentando, em consequência, a satisfação destes clientes com as ofertas de que são alvo.

A aplicação do modelo implicará em campanhas de marketing direto com melhores resultados financeiros, gerando maior lucro para as empresas praticantes. Empresas sadias tendem a empregar mais e manter seus impostos em dia.

6.2. CONCLUSÕES

A Análise RFV Otimizada é uma ferramenta extremamente poderosa para a maximização dos resultados das campanhas de marketing com grande impacto nas estratégias de marketing das empresas que fazem uso do marketing direto. Ao estimar com maior precisão os clientes que irão reagir favoravelmente a uma campanha, a Análise RFV Otimizada aumenta o retorno propiciado pelas campanhas de marketing, aumenta a qualidade da comunicação com os clientes, preserva recursos que podem ser colocados a serviço da organização e de seus clientes. Enfim, torna as organizações mais eficientes e competitivas.

Desta forma, a aplicação deste conceito nas estratégias de marketing das organizações acaba por favorecer a utilização do marketing direto e a conseqüente intensificação do relacionamento entre empresa e cliente.

O RFV Otimizado é uma contribuição inédita para o marketing direto e o marketing de relacionamento, favorecendo um incremento qualitativo e quantitativo nas relações entre as organizações e seus clientes. Este crescimento é qualitativo na medida que otimiza o envio de comunicações para os grupos que apresentam uma resposta positivamente à oferta e, portanto, mostram-se interessados em receber aquela comunicação. No caso da aplicação do RFV Otimizado ao invés da Análise RFV tradicional, estes clientes estão separados de forma otimizada.

O incremento quantitativo está baseado na otimização dos investimentos em produção e comunicação. Se as organizações economizam seus investimentos deixando de comunicar-se com os grupos que não responderão positivamente à oferta, estes investimentos poderão ser revertidos em novas campanhas, com ofertas diferenciadas que serão enviadas aos grupos que irão responder positivamente à nova oferta. Com a aplicação do RFV Otimizado ao invés da Análise RFV tradicional, a quantidade de investimento que é economizada também é otimizada devida à melhor separação entre os clientes que responderão positivamente dos que responderão negativamente.

Por fim, a utilização de um modelo de ajuste dos parâmetros, normalmente arbitrários, da Análise RFV reduz os riscos de execução das campanhas de marketing direto, deixando seus responsáveis mais confortáveis em utilizar esta técnica na estratégia global de marketing da organização.

Portanto podemos resumir que a contribuição fundamental desta tese é mostrar que a despeito de todo o prestígio que a Análise RFV desfruta entre os profissionais de marketing direto, devido principalmente às inúmeras campanhas de marketing cujo sucesso têm sido consistentemente relatado como estando atrelado ao uso do método, ela pode ser melhorada substancialmente quando combinada com o uso de algoritmos genéticos.

Até o momento em que esse trabalho de tese foi finalizado, apesar dos esforços que empreendemos, não tivemos notícia de nenhum outro trabalho que propiciasse um modelo de RFV otimizado. Em conseqüência, também não tivemos notícia de nenhuma outra pesquisa que combine Análise RFV com Algoritmos Genéticos com vistas à otimização do modelo RFV.

6.3. SUGESTÕES DE NOVAS PESQUISAS

O modelo de otimização apresentado nesta tese está baseado nas três dimensões clássicas da Classificação RFV conforme apresentado por Hughes (2000). Porém a este mesmo modelo poderiam ser aplicadas diferentes dimensões de comportamento de compra de clientes. Estas dimensões, que variam de acordo com o mercado, o tipo de produto e o contexto no qual estão sendo utilizadas, também podem ser otimizadas com o modelo e melhorar ainda mais o resultado das campanhas de marketing direto.

Entre as dimensões que poderiam ser aplicadas, está a categoria de produto, conforme sugerido por Schmid (1998). Neste caso, os clientes são classificados também de acordo com a recência, freqüência e valor de cada categoria de produto.

Neste mesmo artigo, Schmid (*op. cit*) sugere também a possibilidade de incluir uma dimensão de sazonalidade que leva em consideração o comportamento anual de compra dos clientes. Alguns clientes são propensos a realizar suas compras em determinados períodos do ano, como por exemplo no natal, no mês de aniversário, no verão, etc. Neste caso os clientes são classificados também de acordo com a porcentagem de vezes que efetuaram compras em determinadas datas. Por exemplo, ao cliente que está cadastrado há cinco anos e que efetuou compras em quatro períodos natalinos atribui-se o índice de sazonalidade 0,8, indicando que efetuou compras em 80% dos períodos natalinos desde o seu cadastramento.

Assim como as decisões relacionadas com as dimensões de recência, frequência e valor são arbitrárias e podem ser otimizadas (número de segmentos, ordem de aplicação das dimensões, etc.), as mesmas decisões precisam ser tomadas com relação às demais dimensões e, da mesma forma, podem sofrer otimização para maximizar o retorno das campanhas de marketing direto.

Uma série de outras características associadas ao comportamento dos clientes no relacionamento com a instituição podem ser testadas pelo processo de otimização. Entre elas citamos: forma de pagamento, canal de comunicação do cliente com a empresa, canal de venda, etc.

Outra característica que poderia ser incluída no modelo do RFV Otimizado é a correção do valor monetário em relação ao tempo. Durante a pesquisa bibliográfica não foi encontrada menção a esta característica em nenhum dos diferentes métodos de Análise RFV, mas ela poderia ser testada para verificar se existe ganho no resultado final da classificação.

Além disso o método proposto por Hughes (*Op. Cit.*) e utilizado nesta tese utiliza apenas uma amostra para realizar o teste. Nos casos em que o resultado da campanha já é conhecido, o modelo poderia ser executado com “n” diferentes amostras retiradas de forma aleatória para que o resultado médio destas execuções fosse o considerado pelo modelo.

REFERÊNCIAS BIBLIOGRÁFICAS

AAKER, D.A, KUMAR, V., DAY, G.S. **Pesquisa de Marketing**. 6 ed. São Paulo: Editora Atlas, 2000. 745p.

ASSOCIAÇÃO BRASILEIRA DE MARKETING DIRETO. **Marketing Direto, o que é?** Disponível na INTERNET via <http://www.abemd.com.br/>. Arquivo consultado em 2001.

ASSOCIAÇÃO BRASILEIRA DE MARKETING DIRETO. **Diretrizes Setoriais**. Disponível na INTERNET via <http://www.abemd.org.br/diretrizes/>. Arquivo consultado em 23/04/2003.

ASSOCIAÇÃO BRASILEIRA DE MARKETING & NEGÓCIOS. **Verbetes**. Disponível na INTERNET via <http://www.abmn.com.br/biblioteca/index.htm>. Arquivo consultado em 23/04/2003.

BANZHAF, W., NORDIN, P., FRANCONI, F. D., KELLER, R. E. **Genetic Programming: An Introduction**. Morgan Kaufmann Publishers, 1998. 470 p.

BAUER, R. J. **Genetic Algorithms and Investment Strategies**. New York: John Wiley & Sons, 1994. 320p.

BITRAN, G., MONDSCHHEIN, S. *A Comparative Analysis of Decision Making Procedures in the Catalog Sales Industry*. **European Management Journal**, Volume 15, Number 2, pp.105-116, 1997.

BLEDSOE, W. W. ***The Use Of Biological Concepts in The Analytical Study Of Systems***. San Francisco: ORSA-TIMS National Meeting, 1961.

BOX, G. E. P. *Evolutionary Operation: A Method for Increasing Industrial Productivity*. **Jounal Of The Royal Statistical Society**, Number 2, pp. 81-101, 1957.

BREMERMANN, H. J. *Optimization Through Evolution and Recombination*. **Self-Organizing Systems**, Spartan Books, 1962.

DARWIN, C. R. ***On The Origin of Species by Means of Natural Selection or The Preservation of Favoured Races in the Struggle for Life***. 1 ed. London: Editor John Murray , 1859. 502p.

DAVIDOR, Y. *Genetic Algorithms And Robotics - A Heuristic Strategy for Optimization*. **Series in Robotics and Automated Systems - Vol. 1**. World Scientific, 1991. 180p.

DAWID, H. ***Adaptive Learning by Genetic Algorithms: Analytical Results and Applications to Economic Models***. 2 ed. Springer Verlag, 1999. 200p.

DIRECT MARKETING ASSOCIATION (DMA). **Honoring Lifetime Archivement: The DMA Hall of Fame Inductees**. Disponível na INTERNET via <http://www.the-dma.org/industryawards/hofinductees.shtml>. Arquivo consultado em 04/06/2003.

EIBEN, A. E., EUVERMAN, T. J., KOWALCZYK, W., PEELEN, E., SLISSER, F., WESSELING, J. A. M. **Comparing Adaptive and Traditional Techniques**

for Direct Marketing. *Proceedings of the 4th European Congress on Intelligent Techniques and Soft Computing*, pp. 434-437. Verlag Mainz. 1996.

EIBEN, A. E., KOUDIJS, A. E., SLISSER, F. **Genetic Modelling of Customer Retention.** *Proceedings of the First European Workshop on Genetic Programming*. Vol. 1391, pp. 178-186, Springer-Verlag, Abril de 1998.

EMARKETER, INC. **The eMail Marketing Report – Executive Summary.** New York, Outubro de 2000. 11p.

EVERITT, B. S., LANDAU, S., LEESE, M. **Cluster Analysis.** Oxford University Press, 4. ed., 2001. 237p.

FITZPATRICK, M. *Statistical Analysis for Direct Marketers – In Plain English.* **Data Management**, Volume 64, Issue 4, pp.54-56, Agosto de 2001.

FOGEL, L. J., OWENS, A. J., WALSH, M. J. **Artificial Intelligence Through Simulated Evolution.** Wiley, 1966.

FRIEDMAN, G. J. *Digital Simulation of an Evolutionary Process.* **General Systems Yearbook**, 4, pp.171-184, 1959.

GEN, M., CHENG, R. **Genetic Algorithms and Engineering Design.** John Wiley & Sons, 1996. 432 p.

GOLDBERG, D. E. **Genetic Algorithms in Search, Optimization, and Machine Learning.** Addison-Wesley, 1989, 372p.

GROSCHE, T., ARMIN, H., ROTHLAUF, F. *A Conceptual Approach for Simultaneous Flight Schedule Construction with Genetic Algorithms. EvoWorkshop 2001*, pp.257-267, Alemanha, 2001.

GUERVÓS, J. J. M. ***Informática evolutiva: Algoritmos Genéticos***. Disponível na INTERNET via <http://geneura.ugr.es/~jmerelo/ie/ags.htm>. Arquivo consultado em 31/07/2003.

GUIMARÃES, M. C., BRISOLA, M. V. ***O ECR Como Ferramenta de Marketing: Da Era da Produção à Era do Valor do Cliente***. Disponível na INTERNET via <http://www.inesc.br/download/O%20marketing%20e%20o%20ECR.pdf>. Arquivo consultado em 30/05/2003.

HAUPT, R. L., HAUPT S. E. ***Practical Genetic Algorithms***. Wiley-Interscience, 1998. 192p.

HOLLAND, J. H. ***Adaptation in Natural and Artificial Systems***. Michigan: University of Michigan Press, 1975.

HUGHES, A. M. ***Strategic database marketing***. New York: McGraw-Hill, 2000. 440p.

IBOPE. **Pesquisas Eleitorais**. Disponível na INTERNET via http://www.ibope.com.br/opp/pesquisa/politica/eleicoes/eleicoes_2002_pres_op_p524_no.htm. Arquivo consultado em 8 de Junho de 2003.

JAIN, L. C., MARTIN, N. M. ***Fusion of Neural Networks, Fuzzy Sets, and Genetic Algorithms: Industrial Applications***. CRC Press, 1998. 354 p.

KOZA, J. R. ***Genetic Programming II: Automatic Discovery of Reusable Programs***. Massachusetts: MIT Press, 1994. 746p.

KOZA, J. R., BENNETT III, F. H., ANDRE D., KEANE, M.A. *Genetic Programming: Biologically Inspired Computation that Creatively Solves Non-Trivial Problems*. **Evolution as Computation**, DIMACS Workshop, pp.15-44, Janeiro de 1999.

KOZA, J. R., KEANE, M. A., STREETER, M. J. *Aperfeiçoando Inventos*. **Scientific American Brasil**, pp.61-67, Março de 2003.

KOTLER, P. ***Administração de Marketing***. 10 ed. São Paulo: Prentice Hall, 2000. 764p.

KWON, Y. K., MOON, B. R. **Personalized Email Marketing with a Genetic Programming Circuit Model**. *Proceeding of the Genetic and Evolutionary Computation Conference (GECCO-2001)*, pp. 1352-1358, Morgan Kaufmann, Julho de 2001.

LEJEUNE, M.A.P.M. *Measuring The Impact Of Data Mining On Churn Management*. **Internet Research: Electronic Networking Applications and Policy**, Volume 11, Number 5, pp.375-387, 2001.

LEWIS, M. A., FAGG, A. H., SOLIDUM, A. *Genetic Programming Approach to the Construction of a Neural Network for Control of a Walking Robot*. **IEEE**

International Conference on Robotics and Automation, pp.2618-2623, França, 1992.

MIGLAUTSCH, J. *Application Of RFM Principles: What To Do With 1-1-1 Customers?* **Journal of Database Marketing**, volume 9, 4, pp.319-324, 2002.

MITCHELL, M. ***An Introduction to Genetic Algorithms***. Massachusetts: MIT Press, 1998. 209p.

MYERS, J. H. ***Segmentation and Positioning for Strategic Marketing Decisions***. Chicago: American Marketing Association, 1996. 358p.

NICOLAI, B. *The 10 Best Catalog Prospecting Ideas*. **Catalog Age**, pp.193-200, Junho de 1999.

ONWUBOLU, G. C. ***Emerging Optimization Techniques In Production Planning And Control***. World Scientific, 2002. 656 p.

PALISADE, CORP. ***Evolver – The Genetic Algorithm Solver for Microsoft Excel***. New York, Agosto de 2001. 205p.

PELTIER, J.W., SCHIBROWSKY, J.A., SCHULTZ, D.E., DAVIS, J. *Interactive Psychographics: Cross-Selling In The Banking Industry*. **Journal Of Advertising Research**, pp.7-22, Março / Abril de 2002.

RAINEY, B. *Transactional Modeling for Beginners*. **Catalog Age**, pp.67, 15 de Março de 2001.

RAPHAEL, M. *Where Did You Come From?* **Direct Marketing**, volume 62, 11, pp.36-38, Garden City, Março de 2002.

RECHENBERG, I. ***Cybernetic Solution Path of an Experimental Problem.*** London: Ministry Of Aviation, Royal Aircraft Establishment, 1965.

REED, J., TOOMBS, R., BARRICELLI, N. A. *Simulation of Biological Evolution and Machine Learning.* **Journal of Theoretical Biology**, 17, pp.319-342, 1967.

ROCHA, A., CHRISTENSEN, C. ***Marketing – Teoria e Prática no Brasil.*** 2 ed. São Paulo: Editora Atlas, 1999. 284p.

ROSENSPAN, A. *The three little pigs: A direct marketing fable.* **Direct Marketing**, volume 59, issue 12, pp.18-21, Abril de 1997.

SCHMID, J., BOYLE, L. *Catalog Creative: The RFMP Way.* **Target Marketing**, pp.38-41, Julho de 1998.

SHETH, J. N., PARVATIYAR, A. *The Evolution of Relationship Marketing.* **International Business Review**, volume 4, number 4, pp.397-418, 1995.

SULLIVAN, J. L. *Optimizing Your Tests.* **Target Marketing**, pp.55-58, Setembro de 2000.

SWENSON, C. A. ***Selling to a Segmented Market.*** Illinois: NTC Business Books, 1996. 177p.

TRIOLA, M. F. ***Introdução à Estatística.*** 7^a. Edição. São Paulo: Editora LTC – Livros Técnicos e Científicos, 1998.

TZAFESTAS, S. G. ***Knowledge-Based Systems - Advanced Concepts, Techniques and Applications***. World Scientific, 1997. 648 p.

WALKENBACH, J. ***Microsoft Excel 2000 Bible***. New York: John Wiley & Sons, Junho de 1999. 936p.

WEDEL, M., KAMAKURA, W. A. ***Market Segmentation – Conceptual and Methodological Foundations***. Boston: Kluwer Academic Publishers, 1998. 378p.

WIENTZEN, H. R. **Abertura da 6ª Conferência Internacional de Database Marketing - CiDBM**. São Paulo, 15 de Maio de 2002.

WINSTON, L. W. ***Financial Models Using Simulation and Optimization: A Step-by-Step Guide With Excel and Palisade's Decisiontools Software***. New York: Palisade Corporation, Junho de 2000. 507p.

YANG, J., HONAVAR, V. ***Feature Subset Selection Using A Genetic Algorithm***. *Artificial Intelligence Research Group, Department of Computer Science, Iowa State University*. Maio de 1997. Disponível na INTERNET via <http://citeseer.nj.nec.com/yang98feature.html>. Arquivo consultado em 05/08/2003.