

Personalização de Serviços de Comércio Eletrônico

Comércio Eletrônico na WWW
Wagner Meira Jr., Cristina Murta e Rodolfo Resende

Personalização

- Adaptação das informações e dos recursos providos aos usuários com a finalidade de incentivar transações comerciais.
- Não afeta a semântica dos serviços, pois aspectos funcionais não são alterados.

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Personalização de Serviços no Café

- Nível Conceitual: inalterado
- Nível de Aplicação: inalterado
- Nível Funcional:
 - Especificação da personalização
 - O que será personalizado?
- Nível de Execução:
 - Detalhamento de todo o processo de personalização

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Nível Funcional

- Para cada serviço:
 - Finalidade
 - facilitar utilização, promover vendas
 - Informações
 - quais informações e procedimentos são personalizados
 - Critérios de personalização
 - preferências dos clientes, produtos correlacionados

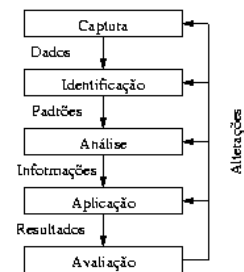
© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Nível de Execução

- Fases:
 - Captura: quais dados são coletados
 - Identificação padrões: determina padrões de comportamento individuais ou coletivos
 - Análise: caracteriza o comportamento
 - Aplicação: personaliza serviços
 - Avaliação: verifica eficácia e determina alterações

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Nível de Execução



© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Captura de Dados

- Fontes de dados:
 - Acessos
 - Referências
 - Agentes
 - Cadastro
 - Informações sobre produtos
- Granulação: sessões

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Representações de Dados

- Expressões regulares
- Hierarquias de conceito
- Regras
- Grafos:
 - CBMG
 - Gramáticas probabilísticas

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Representações de Dados Expressões Regulares

- Templates de navegação
 - Notação:
 - | - Mesma visita
 - * - Qualquer valor com múltiplas ocorrências
 - ; - Entre visitas
 - ^ - Negação
 - ? - Qualquer valor

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Representação de Dados Templates de Navegação

Análise da eficácia de uma campanha de presentes, focando em consumidores que entraram pela página principal, viram o anúncio da campanha e posteriormente voltaram e compraram. Excluir caminhos que incluam visitas às ofertas padrão da loja.

Exemplo:

- (1) <index.htm|*/gifts.htm;*;purchase.htm|?>
- (2) ^<*;reduced.htm;* >
- (3) ^<*;junk.htm;* >

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Representação de Dados Hierarquias de Conceito

- Organizam informações de acordo com critérios semânticos
- Realizada manualmente
- Exemplo:
 - divisão de livros em gêneros
- Utilizadas para correlacionar produtos
 - permitem correlacionar comportamento de usuários

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Representação de Dados Regras

- Indicam o comportamento probabilístico do usuário
- Exemplos:
 - 40% dos usuários que acessam a página principal vão à página de ofertas
- A representação indica ações de caminhamento, sem entrar no mérito semântico.

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Representação de Dados Grafos

- Forma condensada de representar regras de caminhamento.
 - CBMG:
 - Vértices: estados de compra.
 - Arestas: probabilidade de mudança de estado
 - Gramáticas Probabilísticas de Hipertexto:
 - Vértices: páginas
 - Arestas: probabilidade de ir para página

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Descoberta de Padrões

- Técnicas
 - Inteligência artificial
 - Mineração de dados
 - Psicologia
 - Teoria da Informação
- Exemplos:
 - regras de associação
 - padrões sequenciais
 - caminhos mais comuns

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Análise de Padrões

- Determinados os padrões, é necessário:
 - Entender
 - Visualizar
 - Interpretar padrões
- Técnicas
 - OLAP
 - Mecanismos de consulta

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Exemplos de Padrões de Acesso

- Regras de Associação
 - 40% dos clientes que acessaram A também acessaram B
 - 30% dos clientes que acessaram A compraram um produto
- Padrões sequenciais
 - 60% dos clientes que compraram X, também compraram Y no prazo de 15 dias

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Exemplos de Padrões de Acesso

- Classificação e agrupamento
 - clientes que compram A são estudantes
 - clientes que acessam B normalmente visitaram a companhia X
 - 50% dos clientes que compraram C tem entre 20 e 25 anos e moram em BH

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Processo de Descoberta de Padrões de Acesso

- Fases:
 - Limpeza de dados
 - Identificação de transações
 - Integração de dados
 - Transformações
 - Determinação de padrões
 - Análise de padrões

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Fontes de Dados

- Log de Acessos
- Log de Referenciadores
- Log de Agentes
- Dados de cadastro
- Dados de agentes remotos
- Arquivos do site

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Fonte de Dados Log de Acessos

- Campos:
 - IP
 - Data-Hora
 - Requisição
 - Resultado
 - Bytes Enviados
- IP não é suficiente para identificação!
 - IP dinâmico
 - Proxies

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Identificação de Padrões

- Identificação de padrões de utilização de serviços
 - Padrões informativos
 - explicitam informações interessantes
 - agrupamentos, regras de associação
 - Padrões de previsão
 - estimativa de atributo desconhecido com base em dados conhecidos
 - regressão, classificação

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Agrupamentos (*Clusters*)

- Conjuntos cujos elementos são similares em relação a um dado critério.
- Centróide:
 - ponto cujos valores são as médias dos pontos que compõem o cluster
- Algoritmos:
 - k-means
 - árvore geradora mínima

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Algoritmo k-means

Escolha k pontos a serem utilizados como estimativa inicial para os centróides

Repita

Examine cada ponto a ser agrupado e o associe ao centróide mais próximo, atualizando o valor do centróide

Até que nenhum ponto mude de cluster ou o máximo de iterações seja atingido.

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Algoritmo k-means

- Questões
 - função similaridade
 - distância geométrica
 - outras medidas de distância
 - proximidade ortográfica
 - proximidade semântica
 - número de clusters
 - função da finalidade e da natureza dos pontos

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Algoritmo Árvore Geradora Mínima

- Faça cada ponto um cluster
- Enquanto numcluster > objetivo
 - Determine o centróide de cada cluster
 - Calcule a matriz de similaridade entre os clusters
 - Determine os clusters com maior similaridade não máxima (ou seja, idênticos) e faça um novo cluster

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Algoritmo Árvore Geradora Mínima

- Questões
 - Número de clusters ainda é um problema
 - podemos usar distância mínima como critério de parada?
 - Função similaridade
 - Qual é?
 - Qual o seu significado?

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Regras de Associação

Uma implicação da forma $X \Rightarrow Y$, onde X e Y são conjuntos disjuntos de itens de um registro.

Métricas:

- *Suporte*: probabilidade que um registro contenha tanto X quanto Y
- *Confiança*: probabilidade que um registro contenha X , dado que contém Y

Exemplo: análise de dados de supermercados
50% das pessoas que compram X e Y , compram Z

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Geração de Regras de Associação

Definições:

- Seja $I = \{i_1, i_2, i_3, \dots, i_n\}$ um conjunto de itens.
- Seja T um conjunto de transações, onde cada transação t é um conjunto de itens, tal que $t \subseteq I$. Cada transação tem um identificador único (*TID*).
- Sejam $X \subseteq T$ e $Y \subseteq T$ e $X \cap Y = \emptyset$, uma regra de associação é uma implicação da forma $X \Rightarrow Y$.

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Geração de Regras de Associação

Problema:

- Gerar todas as regras que tenham suporte e confiança maiores que os valores mínimos especificados (*minsup* e *minconf*, respectivamente).

Duas etapas:

- Determinar todos os conjuntos de itens cujo *suporte* é maior do que o suporte mínimo especificado. Eles são chamados de conjuntos *freqüentes* de itens.
- Gerar as regras de associação a partir dos conjuntos *freqüentes* de itens. Dados conjuntos *freqüentes* $X Y Z$ e $X Y$, a regra $Z \leftarrow X Y$ é válida se a razão
$$conf = \frac{supp(X Y Z)}{supp(X Y)}$$
é maior que a confiança mínima.

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Geração de Regras de Associação

Problemas:

- **Associabilidade:**
Quais são os itens entre os quais devem ser investigadas as associações.
- **Custo de enumeração:**
Adotar técnicas que reduzam a quantidade de conjuntos a serem enumerados, pois o seu custo computacional cresce exponencialmente com o número de itens.

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Redes Neurais: Recuperação de Padrões

- Mecanismos de mapeamento
 - Autoassociativo: armazena padrões
 - Heteroassociativo: armazena pares de padrões
- Tipos de recuperação
 - Discreta: Padrão mais próximo
 - Contínua: Interpolação

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Redes Neurais

- Sistema Nervoso Humano
 - Terminações sensitivas (entrada)
 - Meissner (tato)
 - Ruffini (calor)
 - Krause (frio)
 - Unidades funcionais (processamento)
 - Neurônios
 - Terminações motoras
 - Músculos
 - Glândulas/Órgãos

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Redes Neurais Artificiais

- Mimetizam as redes naturais em:
 - morfologia e fisiologia dos neurônios
 - padrão de interconexão
 - características excitatórias ou inibitórias
 - terminações sensitivas ou motoras
 - características das sinapses

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Redes naturais X artificiais

- Velocidade de processamento
 - processadores correntes são 10^6 mais rápidos
- Complexidade estrutural
 - 10^{11} a 10^{14} neurônios com 10^3 a 10^4 conexões
- Paralelismo
 - o alto grau de paralelismo das redes naturais permite que problemas como linguagem e visão sejam resolvidos muito mais eficientemente por humanos

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Redes naturais X artificiais

- Armazenamento
 - computador: realocado
 - cérebro: adaptado (ajuste de conexões)
- Tolerância a falhas
 - informações armazenadas de forma difusa
 - o ser humano perde 10^3 neurônios dia
- Controle de processamento
 - não há centralização de controle

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Redes Neurais Artificiais

- Componentes
 - topologia de interconexão de elementos de processamento
 - codificação de dados de entrada
 - recuperação de resultados

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Elementos de Processamento

- Vetor de Entradas $E = (E_1, \dots, E_i, \dots, E_n)$
- Vetor de Pesos $W_j = (W_{1j}, \dots, W_{ij}, \dots, W_{nj})$
- Nível de Ativação T_j
- Função de Ativação f
- Saída S_j

$$S_j = f(E \cdot W_j - T_j) = f(\sum E_i \cdot W_{ij} - T_j)$$

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Função de Ativação

- Mapeia entradas e saídas
 - Linear
 - Rampa
 - Degrau
 - Sigmóide (limitada, monotônica, crescente)
 - Função logística: $1/(1+\exp(-x))$
 - Tangente hiperbólica
 - Razão incrementada de quadrados: $x^2/(1+x^2)$

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Topologia

- Tipos de conexão:
 - Excitatória
 - Inibitória
- Esquema de interconexão
 - Intra-nível
 - Inter-nível
 - Recorrente

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Memória

- Padrões:
 - Espacial: única imagem
 - Espaço-temporal: sequência de padrões espaciais
- Tipos de memória
 - RAM: mapeia endereços para dados
 - CAM: mapeia dados para endereços
 - Memória associativa: mapeia dados para dados

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Aprendizagem

- Qualquer mudança no vetor de pesos
- Tipos:
 - Supervisionado: exemplos externos
 - Espacial: codifica mapeamentos
 - Temporal: codifica sequência de passos
 - Não supervisionado:
 - auto-organiza dados, descobrindo propriedades coletivas

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Aprendizagem por Correção de Erros

- Ajusta os pesos das conexões de acordo com a diferença entre os valores desejados e obtidos em cada nodo da saída
- $D_{wij} = \alpha \cdot O_j \cdot (D_j - O_j)$
- Limitação:
 - só funciona com redes de dois níveis, pois quanto do erro de saída que deve ser atribuído a um nodo intermediário não é definido.

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Back-propagation

- Rede *feed-forward*
 - Não há ciclos
 - Uso para tarefas hetero-associativas
 - Rápida utilização
 - Aprendizagem multi-nível
 - demanda a derivação de uma relação entre a correção no nível de saída e os níveis internos
 - funções de ativação lineares não adicionam graus de liberdade para redes multi-nível

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Back-propagation: produção

- Valor de ativação de cada nodo, dado um padrão de entrada u :
 - $H_{ui} = \sum W_{ji} * E_{uj}$
- Saída do nodo:
 - $O_{ui} = g(H_{ui})$

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Back-propagation: Aprendizagem

- Objetivo: minimizar erro
 - $E = 0.5 * \sum \sum (D_{ui} - O_{ui})^2$
 - Saída: $D_{wji} = B * W_{ji} + A * \sum_u G_{ui} * V_{ui}$
 - V_{ui} = saída do nodo da camada intermediária
 - $G_{ui} = g'(H_{ui}) * (D_{ui} - O_{ui})$
 - Intermediária: $D_{wkj} = B * W_{kj} + A * \sum_u G_{uj} * E_{uk}$
 - E_{uk} = saída do nodo de entrada
 - $G_{uj} = g'(H_{uj}) * \sum_i W_{ji} * G_{ui}$

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Back-propagation

- Questões:
 - parâmetros de treinamento e tamanho da rede devem ser suficientes para comportar o mapeamento desejado
 - natureza de cada entrada e saída deve ser compatível com a capacidade e funcionamento da rede
 - conjunto de treinamento deve garantir cobertura e justiça

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Análise

- Extrair informações dos padrões de comportamento dos clientes
 - Produtos:
 - baseada em aquisições e dados cadastrais
 - Procedimentos:
 - formas prediletas de interação
 - Formato:
 - formatos e informações apresentadas

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Aplicação

- Desenvolver e implementar uma estratégia de personalização a partir da caracterização
- Fatores:
 - mecanismos de personalização
 - escopo
 - forma de atualização

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Formas de Personalização

- Sintetização de páginas
- Ordenação de links
- Regras para notas
- Problemas:
 - Evidências conflitantes
 - Dados poluídos (entrada e regras)

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Regras para notas

- Páginas não podem ser alteradas estáticamente, pois refletem apenas um grupo de usuários.
- Regras para notas são da forma $g, a \rightarrow b$
 - g: padrão de navegação
 - a: página pouco visitada
 - b: página normalmente visitada após g
- Cria um efeito memória

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Mecanismos de Personalização

- Produtos:
 - Venda cruzada: itens complementares
 - Venda ampliada: itens melhores
 - Recomendação: itens de interesse
 - Promoção: itens correlatos em promoção

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Mecanismos de Personalização

- Procedimentos:
 - todos os serviços que recebem parâmetros
 - armazena, para cada usuário, opções e parâmetros utilizados
- Formato:
 - estética: disposição de texto e gráficos
 - conteúdo: tipos de informação a serem apresentados.

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Escopo

- Amplitude do conjunto de usuários a ser personalizado
 - Global: se aplica a todos os clientes ou grupos de clientes
 - Local: se aplica a apenas um cliente
- Compromisso:
 - custo de escopos locais versus personalização insipiente

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Atualização

- Especifica a recentidade dos dados usados para personalização:
 - Discreta: atualização periódica
 - Contínua: atualização instantânea

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Avaliação

- Métricas:
 - taxa de cliques e taxa de conversão
 - taxa de acerto das informações sugeridas
- Micro-métricas:
 - avaliam a efetividade do sítio
 - visita para listagem
 - listagem para cesta
 - cesta para compra

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Privacidade

- Personalização implica invasão de privacidade
 - Insegurança de usuários depende da informação utilizada
- Política de privacidade
 - dados coletados
 - utilização
 - controle de acesso
 - anonimidade

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Jambreiro Busca e Navegação

- Facilitar a localização de produtos
- Funcional:
 - destaca produtos de acordo com os interesses dos clientes, os quais são detectados pelos seus dados cadastrais ou operações de localização e compra anteriores.

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Jambreiro Busca e Navegação

- Execução
 - Captura de dados: registros de acesso, dados cadastrais e compras anteriores
 - Identificação de padrões:
 - segmentar base de produtos: agrupamento
 - correlacionar classes de produtos a dados cadastrais e caminhamento: redes neuronais
 - Análise: determinar um peso para cada grupo de produtos

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Jambreiro Busca e Navegação

- Execução:
 - Aplicação:
 - produtos são enquadrados nas classes
 - classes são ordenadas de acordo com pesos (podem também ser destacados)
 - Avaliação:
 - baseada na efetividade das sugestões, isto é, a sugestão foi aceita

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Jambreiro Colocação na Cesta

- Ampliar a intenção de compra
- Funcional:
 - oferecer produtos correlatos com base em transações anteriores envolvendo o produto sendo colocado na cesta

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Jambreiro Colocação na Cesta

- Execução:
 - Captura de dados:
 - escolhas de usuários em uma mesma sessão ou sessões diferentes, conforme o registro de acessos
 - Identificação de padrões:
 - determinar correlações entre produtos, mais especificamente regras de associação

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Jambreiro Colocação na Cesta

- Execução:
 - Análise: determinar quais regras devem ser usadas para personalização:
 - significância das regras: maior suporte
 - fonte: compras, seleções
 - Aplicação:
 - selecionar produtos para apresentação, com base nas regras mais similares ao caminhar corrente do cliente

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Jambreiro Colocação na Cesta

- Avaliação:
 - baseada em eficácia - quais sugestões foram efetivamente adquiridas, mesmo que em uma visita posterior
 - base de regras pode mudar ao longo do tempo, como consequência do comportamento dos clientes também mudar

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Jambreiro Compra

- Facilitar a interação do cliente com o servidor, minimizando a quantidade de informações a ser provida.
- Funcional:
 - agilizar o processo de compra pelo pré-preenchimento das opções e dados de frete e pagamento.

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Jambreiro Compra

- Execução:
 - Captura de dados:
 - formulários preenchidos anteriormente
 - Identificação de padrões:
 - padrões são simples - valores de campos
 - Análise:
 - selecionar o padrão a ser utilizado, tarefa trivial se houver apenas um.

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.

Jambreiro Compra

- Execução:
 - Aplicação:
 - simples, pré-preenchimento de campos
 - Avaliação:
 - baseada em quantas sugestões são mantidas.

© 2000 Todos os direitos reservados por Meira Jr., Murta e Resende.