**TRABALHO – N2**

GRADUAÇÃO: Tecnologia em Análise e Desenvolvimento de Sistemas

COORDENADOR: Prof. Me. Adriana Ohashi Kei Sato

DISCIPLINA: Estrutura de dados não lineares

DISCENTE: Felipe Schaitel

RM: 24112424

**Valor:** 6,0 pontos

QUESTÕES:

Você é um analista de dados em uma grande empresa de e-commerce. A empresa possui um vasto catálogo de produtos e uma base de clientes diversificada. Recentemente, a equipe de desenvolvimento decidiu melhorar a experiência do usuário implementando uma recomendação de produtos mais eficiente. Para isso, eles planejam utilizar Estruturas de Dados Não-Lineares, como árvores e grafos, para otimizar o processo de recomendação com base no histórico de compras e comportamento de navegação dos clientes.

Desafio:  
A tarefa é desenvolver um sistema de recomendação que possa **sugerir produtos relevantes aos clientes** com base em seu histórico de compras e interações no site. Você deve explicar como utilizaria Estruturas de Dados Não-Lineares para resolver esse problema. Considere os seguintes pontos:

Escolha da Estrutura de Dados:

* Qual estrutura de dados não-linear (árvores, grafos, ou ambas) você escolheria para implementar o sistema de recomendação? Justifique sua escolha. Como a estrutura escolhida pode ajudar a organizar e acessar os dados de maneira eficiente?

Escolheria a estrutura de grafos e tabela hash para implementar no sistema de recomendação.

Grafos são um tipo de estrutura de dados nova, flexível e abstrata. Um grafo é um conjunto de conexões e uma maneira de modelar como diferentes eventos estão conectados entre si, como clientes, produtos etc. Os grafos são constituídos de vértices e arestas. É uma estrutura de dados bastante útil e aplicável em diversos contextos como, por exemplo, sugestão de produtos a partir dos dados do histórico de compras. Isso é possível pois com essa estrutura de dados podemos utilizar o algoritmo de pesquisa em largura. A versatilidade de representação, escalabilidade, modelagem etc dessa estrutura é utilizada em diversos setores da indústria etc.

Enquanto a tabela hash usa uma função hash para indicar, de maneira inteligente, onde armazenar os elementos. Elas também são conhecidas como mapas hash, mapas, dicionários e tabelas de dispersão. Além disso, as tabelas hash são muito rápidas! As tabelas hash funcionam bem com grandes volumes de dados além de que, geralmente, as linguagens de programação já possuem tabela hash implementado. A tabela hash será utilizada para armazenar eficientemente os dados que precisam ser acessados rapidamente como o histórico do usuário e relacioná-lo a cada item que já interagiu, comprou, gostou, pesquisou etc.

Modelagem do Problema:

* Descreva como modelaria os dados do histórico de compras e das interações dos clientes usando a estrutura de dados escolhida. Quais seriam os principais nós e arestas (ou ramos) nesta estrutura?

A modelagem do sistema seria feita da seguinte maneira: cada usuário seria representado pelo vértice (nó) e os produtos são o outro conjunto de vértices (nós). As arestas que ligam e relacionam os vértices seriam referentes às interações como:

cliente ---> produto

produto ---> produto

As tabelas hashs são suportes para acessar os dados relacionados aos grafos através das chaves e valores.

Algoritmos de Busca e Recomendação:

* Explique quais algoritmos de busca você utilizaria para percorrer a estrutura de dados e encontrar produtos relevantes para recomendação.

Grafos é eficiente para a funcionalidade de recomendação pois com essa estrutura de dados podemos utilizar o algoritmo de pesquisa em largura. A pesquisa em largura permite encontrar o menor caminho entre 2 ou mais pontos/produtos/objetos. Esse algoritmo resolve e busca o caminho mínimo para algo. Assim, encontra o item mais próximo de você. Também é possível aplicar para buscar como:, um mercado consumidor, fornecedores etc. Esse algoritmo estabelece e analisa também os graus de conexão. Sempre sendo priorizados os de primeiro grau pois é o caminho com menos etapas, o caminho mínimo.

* Como garantiria que as recomendações são personalizadas e relevantes para cada cliente?

Ouvindo diretamente os usuários e *stakeholders*. Inserir na interface um espaço para ser preenchido com sim ou não se o produto sugerido é relevante.

Desafios e Soluções:

* Quais desafios você prevê ao implementar este sistema utilizando estruturas de dados não-lineares?

A implementação e a manutenção da estrutura de grafos é complexa. Grafos que não foram bem estruturados demandam espaços a mais de armazenamento e isso influencia negativamente no processamento tornando-o mais lento, além de ter o custo mais elevado quando há grafos de grandes dimensões pois vai exigir mais recursos computacionais.

Nas tabelas hashs os valores (números e strings) são armazenados em chaves. Entretanto, pode acontecer de duas chaves serem indicadas para o mesmo espaço/indice ocasionando uma colisão.

* Proponha possíveis soluções ou alternativas para esses desafios.

A solução para esse problema está em investir tempo e recursos para realizar uma pesquisa profunda referente a análise de requisitos. A fim de documentar, definir as funcionalidades e o tipo de banco de dados de maneira completa. Todos esses elementos influenciam a boa modelagem do sistema e consequentemente a estrutura do grafo.

Uma solução alternativa para o problema de colisão seria a vista no livro “Aprendendo algoritimos de Forma Ilustrada” pois segundo Bhargava na página 141 “Colisões são um problema, e você precisa solucioná-lo. Para isso há várias alternativas, e a mais simples é esta: se diversas chaves mapeiam para o mesmo espaço, inicie uma lista encadeada neste espaço”.

Referências Bibliográficas:

* BHARGAVA, Aditya Y. Aprendendo algoritmos de forma ilustrada. Trad. [BrodTec]. 1. ed. [São Paulo]: [Novatec], [2017].