**Pandey2020\_Article\_HowGoodAreModernSpatialLibrari**

Em que área do conhecimento o artigo se encaixa? Qual o contexto geral da utilização/aplicação dessa pesquisa? Qual é o problema?

Muitos aplicativos e serviços dependem da localização de seus usuários como: Uber,Tinder, Ifood etc. Para conseguirem tais dados constroem seus próprios sistemas de gerenciamento de dados espaciais ou utilizam soluções existentes. O artigo compara as bibliotecas mais utilizadas e mostra a eficiência de cada visto que são amplamente utilizadas e devido a variedade de soluções as empresas podem estar utilizando bibliotecas sem eficiência para seu modelo de negócio. O problema da utilização de bibliotecas de geometria planares diretamente com coordenadas GSP.

**Trabalhos relacionados**

a. Soluções existentes: quais os artigos que já investigaram esse problema?

b. Solução mais próxima: qual é a melhor e/ou solução mais próxima ao problema apresentado?

a. Comparative evaluation of various indexing techniques of geospatial vector data for processing in distributed computing environment, Large-scale spatial join

query processing in cloud.

b. Cada biblioteca tem seus pontos positivos e negativos, a escolha fica a critério do usuário que vai utilizar e do problema a ser resolvido. O problema está nas aplicações calcular a distância euclidiana entre duas coordenadas de GPS brutas.

**Gap**

Qual a principal/is limitações da(s) solução(ões) existente(s)? O que não foi investigado? Quais aspectos do problema ainda não estão claros? Por que as abordagens existentes podem ser falhas?

Cada biblioteca atende um tipo de solução e além das existentes as empresas constroem suas próprias soluções para atenderem seu modelo de negócio, com isso não é possível determinar que aquele é menos eficiente ou que é a melhor. Não foi investigado dados do tipo linestring, polylines. Não existe biblioteca perfeita que atenda todas as demandas das aplicações e serviços, além de muitas estarem em constante desenvolvimento, com isso a avaliação feita hoje pode mudar amanhã devido a melhorias.

**Hipótese ou questão de pesquisa**

Dado a limitação / problema listado anteriormente, o autor argumenta que há uma forma de fazer diferente que talvez resolva o problema, ou ainda atinja resultado melhor ou que investigue algum aspecto?

A abordagem correta é projetar as coordenadas GPS brutas utilizando um sistema de referência espacial.

**Objetivo**

Mostrar os problemas de usar bibliotecas de geometria plana com coordenadas GPS e avaliar o desempenho de diferentes tipos de bibliotecas.

**Metodologia:**

a. Qual o tipo de revisão realizada: narrativa, sistemática, integrativa, ou outro tipo?

Foi utilizado a revisão sistemática

b. Quais são as amostras, as variáveis, tipo de pesquisa e as métricas utilizadas?

Os conjuntos de dados utilizados foram: NYC Taxi and Limousine Commission (TLC)—TLC Trip Record

Data (2019) e geo-tagged tweets in the New York City area (NYC

Tweets) utilizando API do Twitter para coletar os dados. Para medir o desempenho utilizaram ferramentas de benchmark e parâmetros específicos nas consultas.

c. Como foi realizada a coleta e análise de dados?

Foi utilizado dados públicos fornecidos pela cidade de NYC e alguns dados coletados pelo Twitter utilizando API. As análises foram com bases nas consultas espaciais de range, distance, k-NN, and a spatial join query.

**Machine Learning for Crowdsourced**

**Spatial Data**

Em que área do conhecimento o artigo se encaixa? Qual o contexto geral da utilização/aplicação dessa pesquisa? Qual é o problema?

Dados espaciais vem cada veze mais sendo utilizados, porém criar e manter um banco de dados espaciais não é fácil e tem um custo alto. Os mapas de crowdsourced tem um potencial para fornecer dados do nosso mundo, porém ainda se tem dúvidas sobre a qualidade dos dados.

**Trabalhos relacionados**

a. Soluções existentes: quais os artigos que já investigaram esse problema?

b. Solução mais próxima: qual é a melhor e/ou solução mais próxima ao problema apresentado?

a. How good is volunteered geographical information? a comparative study

of openstreetmap and ordnance survey datasets.

b. O artigo analisado apresenta machine learning como forma de medir e melhorar a qualidade dos dados. Cada vez mais surge novas técnicas e ferramentas que nos auxiliam a criar coisas novas, machine learning tem um grande potencial e acredito que utilizar tecnologias como essa para medir e melhorar o que já existe seja uma alternativa melhor do que comparar com soluções que já existe. Foi utilizado a acurácia como métrica.

**Gap**

Qual a principal/is limitações da(s) solução(ões) existente(s)? O que não foi investigado? Quais aspectos do problema ainda não estão claros? Por que as abordagens existentes podem ser falhas?

As abordagens existentes avaliam a qualidade por meio de mapas oficiais que já existem, ou seja, comparar algo que já existe com algo novo pode parar novas.

**Hipótese ou questão de pesquisa**

Dado a limitação / problema listado anteriormente, o autor argumenta que há uma forma de fazer diferente que talvez resolva o problema, ou ainda atinja resultado melhor ou que investigue algum aspecto?

O autor ressalta que comparar os dados OSM com outras formas de mapas como: GoogleMaps é ineficiente para um melhor resultado sobre a qualidade dos dados propõem o uso de machine learning.

**Objetivo**

O autor vai utilizar machine learning para medir a qualidade dos dados OSM.

**Metodologia:**

a. Qual o tipo de revisão realizada: narrativa, sistemática, integrativa, ou outro tipo?

Foi utilizado a revisão sistemática

b. Quais são as amostras, as variáveis, tipo de pesquisa e as métricas utilizadas?

As amostras foram fornecidas do banco de dados de Londres e banco de dados OSM.

c. Como foi realizada a coleta e análise de dados?

Foi utilizado dados abertos fornecido por Londres e da própria solução.