

Georeferenciamento e Bancos de Dados NoSQL: Uma análise do panorama atual na agropecuária de precisão

Ricardo Peixoto Robaina, Sandro da Silva Camargo

Programa de Pós-Graduação em Computação Aplicada
Universidade Federal do Pampa & Embrapa Pecuária Sul
Bagé, Rio Grande do Sul, Brasil

ricardorobaina11@gmail.com, sandro.camargo@unipampa.edu.br

Resumo O recente emprego de tecnologias de informação nos sistemas de produção agrícolas, visando uma maior produtividade, traz desafios associados. O enorme volume de dados heterogêneos gerados continuamente leva a um cenário de *Big Data*, no qual sistemas de bancos de dados tradicionais padrão SQL, baseados em esquemas, apresentam algumas deficiências. As tecnologias de bancos de dados do padrão NoSQL são uma alternativa promissora nesse contexto. Dentre estes, o banco de dados MongoDB é um dos citados em aplicações de Agropecuária de Precisão. O presente artigo apresenta um estudo acerca do estado da arte da utilização do banco de dados MongoDB, na Agropecuária de Precisão, bem como destaca as suas principais características com potencial utilização nesse cenário.

Palavras-chave: MongoDB, Big Data, Sistemas de Informação.

1 Introdução

Nos últimos 50 anos, houve a duplicação população do planeta. Além disso, estimativas apontam que no ano de 2050, a população mundial deve se aproximar de 10 bilhões de pessoas. Esse fato intensifica a demanda pela produção de alimentos. Devido à quantidade finita de recursos naturais em nosso planeta, o aumento da produtividade em sistemas de produção é essencial [13].

A Agropecuária de Precisão (AP) pode contribuir para o aumento da produtividade através do emprego de dispositivos computacionais no monitoramento do processo produtivo nas tarefas de coleta, armazenamento, processamento e análise de grandes quantidades de dados, com o objetivo de produzir conhecimento útil e auxiliar os processos de tomada de decisão do produtor rural [9]. A coleta automatizada gera continuamente um grande volume de dados heterogêneos que necessitam ser armazenados, caracterizando um cenário de *Big Data* [14]. Karmas e colaboradores [10] apontam que nesses casos os bancos de dados padrão *Not only SQL* (NoSQL) são mais adequados em comparação

aos bancos de dados relacionais (SQL). Dentre os bancos de dados NoSQL, o MongoDB é um dos mais utilizados [9].

O objetivo do presente trabalho é investigar a utilização do banco de dados MongoDB em cenários de Agropecuária de Precisão (AP), bem como identificar as funcionalidades dessa ferramenta que possam ser exploradas nesse contexto. O texto está organizado da seguinte forma: a Seção 2 apresenta a metodologia de trabalho executada nesse estudo; em seguida, a Seção 3 aborda os resultados obtidos nessa pesquisa. Por fim, na última são apresentadas as conclusões alcançadas no escopo dessa pesquisa.

2 Metodologia

A metodologia de pesquisa do presente trabalho foi dividida em quatro etapas principais. A Figura 1 ilustra, em um diagrama de seqüência, as etapas definidas nesta metodologia.

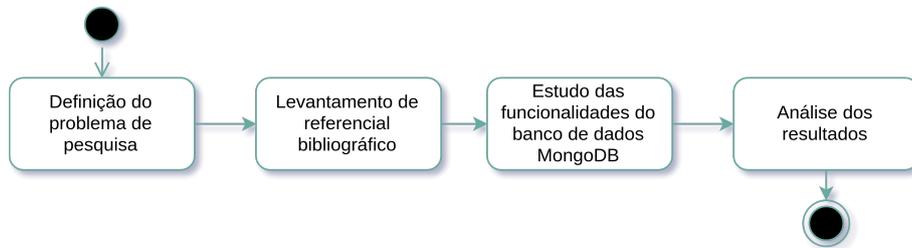


Figura 1: Diagrama de seqüências das etapas da metodologia

O objetivo da primeira etapa dessa metodologia foi definir o problema de pesquisa a ser abordado no presente estudo. Em conjunto com as etapas de coleta e tratamento, o armazenamento de dados é um componente fundamental para as ferramentas de Agropecuária de Precisão [9], e alguns autores apontam as falhas dos BD relacionais neste contexto [2,3]. Além disso, os aspectos espacial e temporal dos dados são dois componentes fundamentais em vários sistemas de agropecuária de precisão [3,9].

O processo de levantamento de referencial bibliográfico foi realizado, através de uma revisão sistemática da literatura, durante a segunda etapa. Definiu-se como fonte de busca o periódico *Computers and Electronics in Agriculture* [5] por abordar especificamente os avanços do desenvolvimento de aplicações computacionais de *hardware* e *software* em indústrias relacionadas à Agropecuária e por possuir um fator de impacto igual a 2.427. As palavras-chave definidas foram as seguintes: MongoDB, NoSQL e georreferenciamento. A consulta retornou um total de cinco documentos que foram identificados como relevantes no escopo desse trabalho.

A terceira etapa destinou-se ao estudo dos artigos elicitados na etapa anterior, bem como a análise das funcionalidades do banco de dados MongoDB. Para isso, foi usada como fonte de estudo a documentação oficial da ferramenta [11]. Utilizou-se dados georreferenciados de estações meteorológicas do Instituto Nacional de Meteorologia do Brasil (INMET) [8] nos testes realizados.

Por fim, na última etapa, fez-se uma análise acerca dos conhecimentos adquiridos nas etapas anteriores.

3 Resultados e Discussão

3.1 Trabalhos Correlatos

Bimonte e colaboradores [3] apresentam um Sistema de Informação de Gestão Agrícola, do inglês *Farm Management Information Systems* (FMIS), chamado *VBoxReporting*. O sistema apresentado é voltado à análise de consumo energético em implementos agrícolas, para apoio à tomada de decisão. O sistema coleta dados de localização e de nível de combustível através de sensores de baixo custo instalados nos equipamentos. Nesse trabalho, os dados georreferenciados são armazenados no banco de dados PostGIS e a visualização dos dados é gerada pelo FMIS. Os autores salientam a necessidade de migração do sistema do banco de dados atual para um banco de dados do padrão NoSQL, a fim de garantir fatores como escalabilidade e performance. Como uma das perspectivas futuras, os autores estão investigando a possibilidade de migração do banco de dados PostGIS para o MongoDB.

Perondi e colaboradores [12] apresentam o processo de desenvolvimento de uma ferramenta de suporte à decisão para escolha de época de semeadura. A demanda por tal ferramenta é oriunda da necessidade de reduzir o risco de ocorrência de eventos climáticos extremos durante fases críticas do crescimento da planta. Para o correto funcionamento do sistema, o usuário deve selecionar manualmente a estação meteorológica mais próxima à sua propriedade, além do cultivo que deseja plantar. Com estes dados, o sistema executa diferentes simulações com o intuito de indicar a probabilidade da ocorrência de eventos climáticos indesejados. A ferramenta desenvolvida no escopo dessa pesquisa foi inserida no projeto AgroClimate [1], o qual disponibiliza diferentes ferramentas online de apoio à decisão para produtores rurais nos EUA. O projeto é mantido em uma infraestrutura Linux e utiliza o banco de dados MongoDB. A base de dados é composta por dados de colheita, cultivo, dados meteorológicos e informações das estações, além de configurações do sistema. Os autores afirmam que a escolha do banco de dados foi embasada no fato da tecnologia ser flexível e suportar uma alta taxa de transferência de dados.

Baker e colaboradores [2] incorporam o banco de dados MongoDB em um sistema de predição da ocorrência de doenças no cultivo de batata. O sistema utiliza dados de previsão meteorológica do *National Weather Service* e gera mapas de calor. Devido à limitação de tamanho máximo de 16 megabyte (Mb) por documento armazenado no banco de dados, fez-se necessário o uso da funcionalidade GridFS, do próprio MongoDB. Tal funcionalidade divide documentos

maiores que o limite máximo em documentos menores para efetuar o armazenamento. Os autores afirmam que a escolha da utilização do MongoDB foi acertada por ter uma menor complexidade de configuração e de uso, em comparação ao banco de dados PostgreSQL, e por suportar dados georreferenciados. Além disso, os autores identificaram os seguintes fatores durante a execução do MongoDB: alto consumo de memória, alta velocidade nas operações e baixa utilização de CPU.

Com o objetivo de avaliar o efeito de ondas eletromagnéticas (Wi-Fi) em colônias de abelhas, Henry e colaboradores [6] utilizam o banco de dados MongoDB para armazenamento de dados de temperatura, umidade, amplitude e frequência sonora no interior de colmeias. Os dados armazenados são disponibilizados em tempo real, através de gráficos gerados em uma aplicação web. Este é um exemplo do emprego do MongoDB no armazenamento e na disponibilização de diferentes tipos de dados.

Trabalho	[3]	[12]	[2]	[6]
Contexto	Análise de consumo energético.	Predição de época de semeadura.	Predição da ocorrência de doenças.	Efeito do Wi-Fi em colmeias.
Banco de Dados	PostGIS	MongoDB	MongoDB	MongoDB
Dados georreferenciados	X	X	X	
Operações georreferenciadas				
Satisfação com banco de dados atual		X	X	X

Tabela 1: Comparação entre os trabalhos correlatos

Os artigos avaliados no escopo desse trabalho apresentam diferentes utilizações do banco de dados MongoDB na Agropecuária de Precisão. A Tabela 1 agrupa informações relevantes entre os trabalhos avaliados. O MongoDB possui suporte a diferentes tipos de dados georreferenciados e a ferramenta MongoDB Compass oferece funcionalidades de visualização de dados adequadas para análises simples como a apresentada por Bimonte, Naofal e Ginesteem [3]. Neste contexto, nota-se a subutilização das potencialidades do banco de dados apresentado. No sistema desenvolvido por Perondi e colaboradores [12], o usuário deve escolher manualmente a estação meteorológica mais próxima. O MongoDB possui operadores de consulta georreferenciados que retornam um ponto mais próximo de outro. Dessa forma, poderia ser facilmente implementada a funcionalidade de escolha automática de estação. Baker e colaboradores [2] apresentam vantagens da utilização do banco de dados MongoDB em comparação ao banco de dados PostgreSQL. Esse trabalho também apresenta uma solução para a limitação de tamanho máximo de documento. Estes pontos ressaltam que este banco de dados pode ser utilizado inclusive com arquivos maiores que 16 megabytes, como imagens e vídeos com alta resolução. Henry e colaboradores [6] exemplificam a utilização do MongoDB para um cenário de coleta constante de dados gerados por diferentes sensores. A flexibilidade e o desempenho apresentados pelo banco de dados, nesse contexto, indicam o potencial de uso do MongoDB em cenários de IoT (*Internet of Things*).

3.2 Funcionalidades do banco de dados MongoDB

MongoDB é um banco de dados, do padrão NoSQL, orientado a documentos. Diferentemente dos bancos de dados relacionais, nos quais os dados são armazenados em tabelas, nesse tipo de banco, os dados são armazenados em documentos e estes são agrupados em coleções [11]. O formato *JavaScript Object Notation* (JSON) [4] é utilizado como padrão nos documentos do MongoDB.

Outra diferença é que, em bancos de dados relacionais, as tabelas respeitam uma estrutura pré-definida. Já os documentos do MongoDB são flexíveis, uma vez que cada documento tem o seu próprio conjunto de atributos, independente dos outros documentos da mesma coleção. Esse fato colabora para a escalabilidade de sistemas que utilizam esse banco de dados.

Ademais, o MongoDB possui suporte a dados georreferenciados [11]. O padrão GeoJSON [7] é utilizado pra especificar o tipo de dado espacial tratado. Esse formato possui suporte a três tipos principais de dados: ponto, linha e polígono. Além de dados específicos, o banco de dados MongoDB possui operadores de consultas georreferenciadas. Entre os principais operadores de consulta, pode-se citar: *geoIntersects*, o qual retorna todos objetos que intersectam uma geometria; *geoWithin*, que retorna todos objetos no interior de raio determinado; e, *near*, o qual retorna o objeto mais próximo de um ponto.

A Figura 2 apresenta a visualização de uma consulta georreferenciada, gerada pela ferramenta MongoDB Compass. Nessa figura, as estações meteorológicas do Instituto Nacional de Meteorologia (INMET) [8] são representadas por pontos na cor azul. O círculo na cor laranja representa uma consulta por todas as estações que estão dentro de um determinado raio, a partir de um ponto de origem.

Com relação a ferramentas, o MongoDB disponibiliza quatro principais: MongoDB Shell, MongoDB Compass, MongoDB Atlas e MongoDB Mobile. A primeira delas é uma interface por linha de comando para realizar operações em bases de dados. MongoDB Compass é a interface gráfica oficial para acesso a bases de dados, além disso, produz visualizações de dados adequadas de acordo com o tipo de dado armazenado. A terceira, por sua vez, é a plataforma de *cloud computing* do MongoDB, na qual é possível criar e acessar facilmente bases de dados remotamente. Ressalta-se a possibilidade de criar base de dados gratuitamente para a realização de estudos. Por fim, MongoDB Mobile é uma versão do MongoDB específica para dispositivos móveis como *smartphones* e dispositivos de IoT. Essa versão possui uma funcionalidade de sincronização de dados, isto é, armazena dados localmente quando não existe conexão com a internet e sincroniza os dados remotos assim que uma nova conexão for estabelecida.

4 Considerações finais

A maximização da produtividade buscada pela Agropecuária de Precisão contribui para um cenário de geração de um grande volume de dados heterogêneos. Nesse cenário de *Big Data*, os bancos de dados do padrão NoSQL apresentam-se mais adequados por possuírem uma maior flexibilidade e escalabilidade. O

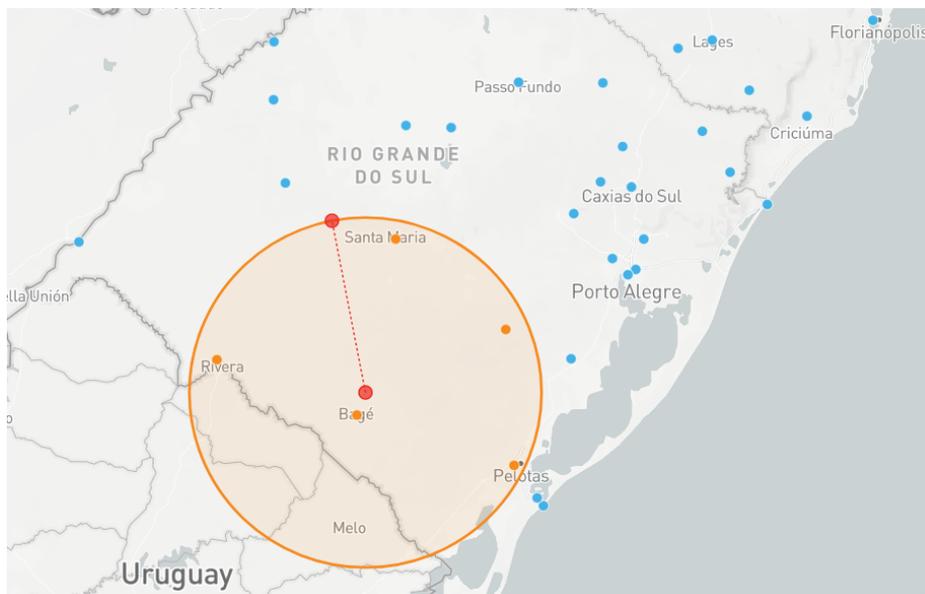


Figura 2: Representação gráfica de uma consulta georreferenciada

banco de dados MongoDB apresenta-se como uma ferramenta com potencialidades necessárias para aplicabilidade na Agropecuária de Precisão. Isso se dá pela facilidade de uso, compatibilidade com dispositivos móveis e suporte a dados e operações georreferenciados. Apesar da diversidade de funcionalidades identificadas no MongoDB, nota-se uma sub-utilização dessa ferramenta na Agropecuária de Precisão, como operações de consultas sobre dados georreferenciados, utilização do MongoDB Mobile em dispositivos de IoT e análises de dados a partir da visualização de dados gerada pela ferramenta MongoDB Compass. Em contrapartida, percebe-se que, apesar de ainda pouco utilizado, o MongoDB tem aumentado sua participação em sistemas de Agropecuária de Precisão. Conclui-se que existe um potencial de exploração das funcionalidades no banco de dados MongoDB no contexto alvo deste estudo.

Agradecimentos: Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de mestrado.

Referências

1. Agroclimate, <http://agroclimate.org/>
2. Baker, K., Roehsner, P., Lake, T., Rivet, D., Benston, S., Bommersbach, B., Kirk, W.: Point-trained models in a grid environment: Transforming a potato late blight risk forecast for use with the national digital forecast database. *Computers and electronics in agriculture* 105, 1–8 (2014)

3. Bimonte, S., Naoufal, E., Gineste, L.: A system for the rapid design and implementation of personalized agricultural key performance indicators issued from sensor data. *Computers and Electronics in Agriculture* 130, 1–12 (2016)
4. ECMA: Standard ecma-404: The json data interchange syntax (2017), <https://www.json.org/>
5. Elsevier: Computers and electronics in agriculture, <https://www.journals.elsevier.com/computers-and-electronics-in-agriculture>
6. Henry, E., Adamchuk, V., Stanhope, T., Buddle, C., Rindlaub, N.: Precision apiculture: Development of a wireless sensor network for honeybee hives. *Computers and Electronics in Agriculture* 156, 138–144 (2019)
7. (IETF), I.E.T.F.: The gejson format (2016), <https://tools.ietf.org/html/rfc7946>
8. INMET: Normas climatológicas do brasil (1981 - 2010), <http://www.inmet.gov.br/portal/index.php?r=clima/normaisClimatologicas>
9. Kamilaris, A., Kartakoullis, A., Prenafeta-Boldú, F.X.: A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture* 143, 23–37 (2017)
10. Karmas, A., Tzotsos, A., Karantzalos, K.: Geospatial big data for environmental and agricultural applications. In: *Big Data Concepts, Theories, and Applications*, pp. 353–390. Springer (2016)
11. MongoDB: Mongodb documentation, <https://docs.mongodb.com/>
12. Perondi, D., Fraisse, C.W., Staub, C.G., Cerbaro, V.A., Barreto, D.D., Pequeno, D.N., Mulvaney, M.J., Troy, P., Pavan, W.: Crop season planning tool: Adjusting sowing decisions to reduce the risk of extreme weather events. *Computers and Electronics in Agriculture* 156, 62–70 (2019)
13. Saath, K.C.d.O., Fachinello, A.L.: Crescimento da demanda mundial de alimentos e restrições do fator terra no brasil. *Revista de Economia e Sociologia Rural* 56(2), 195–212 (2018)
14. Wolfert, S., Ge, L., Verdouw, C., Bogaardt, M.J.: Big data in smart farming—a review. *Agricultural Systems* 153, 69–80 (2017)