



BIG DATA

FOR SOCIAL GOOD

Estudo de caso: Telefónica Brazil

Prevendo níveis de poluição do ar com 24 a 48 horas de antecedência em São Paulo, Brasil

Telefónica

LUCA
Telefónica DATA UNIT

vivo

A GSMA representa os interesses das operadoras móveis de todo o mundo, reunindo quase 800 operadoras e mais de 300 empresas do mais amplo ecossistema móvel, incluindo fabricantes de telefones e dispositivos móveis, empresas de software, fornecedores de equipamentos e companhias de internet, além de organizações de setores industriais adjacentes.

A iniciativa Big Data for Social Good reúne organizações públicas e privadas para acelerar a influência das indústrias de telefones móveis em relação aos ODS da ONU. Doenças infecciosas, poluição, terremotos, enchentes e outros desastres estão entre os maiores desafios que o mundo enfrenta hoje em dia. De acordo com o Banco Mundial, a poluição do ar emergiu como o quarto principal fator de risco para mortes em todo o mundo.

Operadoras móveis podem fornecer insights importantes e únicos com base em dados de rede agregados e anônimos para ajudar a resolver esses problemas

complexos. Os dados móveis podem ajudar as organizações de saúde pública a responderem de forma mais eficaz às epidemias ou planejar intervenções de saúde direcionadas. Pode dar apoio às organizações que prestam serviços em caso de emergência a direcionarem seus recursos de maneira mais precisa e eficiente. Pode ajudar os governos a entenderem melhor o impacto da poluição e das mudanças climáticas para os cidadãos.

Por meio da GSMA, operadoras móveis de vários lugares do mundo se reuniram para acelerar e escalar a oportunidade para o Big Data for Social Good. A GSMA oferece uma plataforma única para estabelecer uma estrutura comum e abordagens baseadas em melhores práticas, respeitando e protegendo a privacidade dos indivíduos. Este estudo de caso do projeto foi desenvolvido como parte da iniciativa da GSMA "Big Data for Social Good", que busca difundir modelos semelhantes que estão sendo implementados em outras grandes cidades do mundo.

1. <http://www.worldbank.org/en/news/infographic/2016/09/08/death-in-the-air-air-pollution-costs-money-and-lives>

Resumo

A Telefónica Brasil está trabalhando com a cidade de São Paulo para aproveitar os dados de rede móvel para combater o impacto negativo da poluição do ar na saúde da população. A Telefónica fornece dados móveis anônimos, algoritmos e ferramentas, que podem ser integrados pelos municípios de São Paulo em seus processos de gerenciamento de trânsito e contaminação do ar. Desenvolvidos pela equipe da LUCA Data Driven Decisions, os algoritmos Telefónica utilizam a *machine learning* e dados anônimos da rede móvel, combinados com dados de sensores climáticos, de trânsito e de poluição, para mapear e prever níveis de poluição em toda a cidade.

A solução pode prever, com antecedência, níveis de poluição de 24 a 48 horas, possibilitando às autoridades locais em São Paulo a tomarem medidas preventivas,

caso as emissões de dióxido de nitrogênio (NO₂) comprometam a saúde humana. Ao empregar big data com dados móveis, a Telefónica e a cidade podem, significativamente, reduzir os custos, aumentar a precisão e fornecer informações importantes aos responsáveis sobre como definir as políticas de transporte urbano ou avaliar os níveis de poluição.

A prova de conceito realizada com a base de dados de São Paulo mostrou que big data com dados móveis é uma ferramenta poderosa para prever e analisar padrões de trânsito e a resultante poluição do ar. Dados móveis anônimos possibilitam um mapeamento mais granular e econômico da contaminação do ar na cidade em comparação a uma rede de sensor estático que coleta dados equivalentes.

O desafio da poluição do ar

Vários estudos mostraram a relação entre a poluição ambiental e problemas de saúde pública. A Organização Mundial da Saúde (OMS) constatou que a poluição do ar está associada a doenças crônicas e que pode agravar patologias relacionadas, como morbidade e mortalidade cardiovascular e respiratória². Um estudo no Brasil³ estimou que mais de 20 mil mortes por ano, em 29 regiões metropolitanas brasileiras, podem estar atribuídas à poluição do ar.

Como há mais dados ambientais cada vez mais disponíveis, há uma tendência mundial em relação a utilizar abordagens baseadas em dados para gerenciar a poluição do ar e seu impacto, em consonância com os Objetivos de Desenvolvimento Sustentável (ODS) estabelecidos pelas Nações Unidas:

- **ODS 3** Assegurar uma vida saudável e promover o bem-estar para todas e todos, em todas as idades: o ODS 3 busca aumentar a expectativa de vida, em parte por meio de medidas para reforçar a capacidade dos países, particularmente os países em desenvolvimento, para o alerta precoce, redução de riscos e gerenciamento de riscos nacionais e globais de saúde.
- **ODS 13** Tomar medidas urgentes para combater a mudança climática e seus impactos.
- **ODS 11** Tornar as cidades e os assentamentos humanos inclusivos, seguros, resilientes e sustentáveis: O ODS 11 procura reduzir o impacto ambiental negativo per capita das cidades, até 2030, prestando especial atenção à qualidade do ar e gestão de resíduos.

O mapeamento preciso dos níveis de poluição em áreas metropolitanas pode ser usado para:

- Identificar os níveis potenciais de poluição local que possam afetar a saúde da população local e, a longo prazo, vincular essas informações a outros dados populacionais (por exemplo, asma) para informar proativamente os cidadãos sobre o impacto específico.
- Informar os sistemas de gestão de trânsito a reduzir o número de veículos em áreas com níveis altos ou perigosos de poluição.

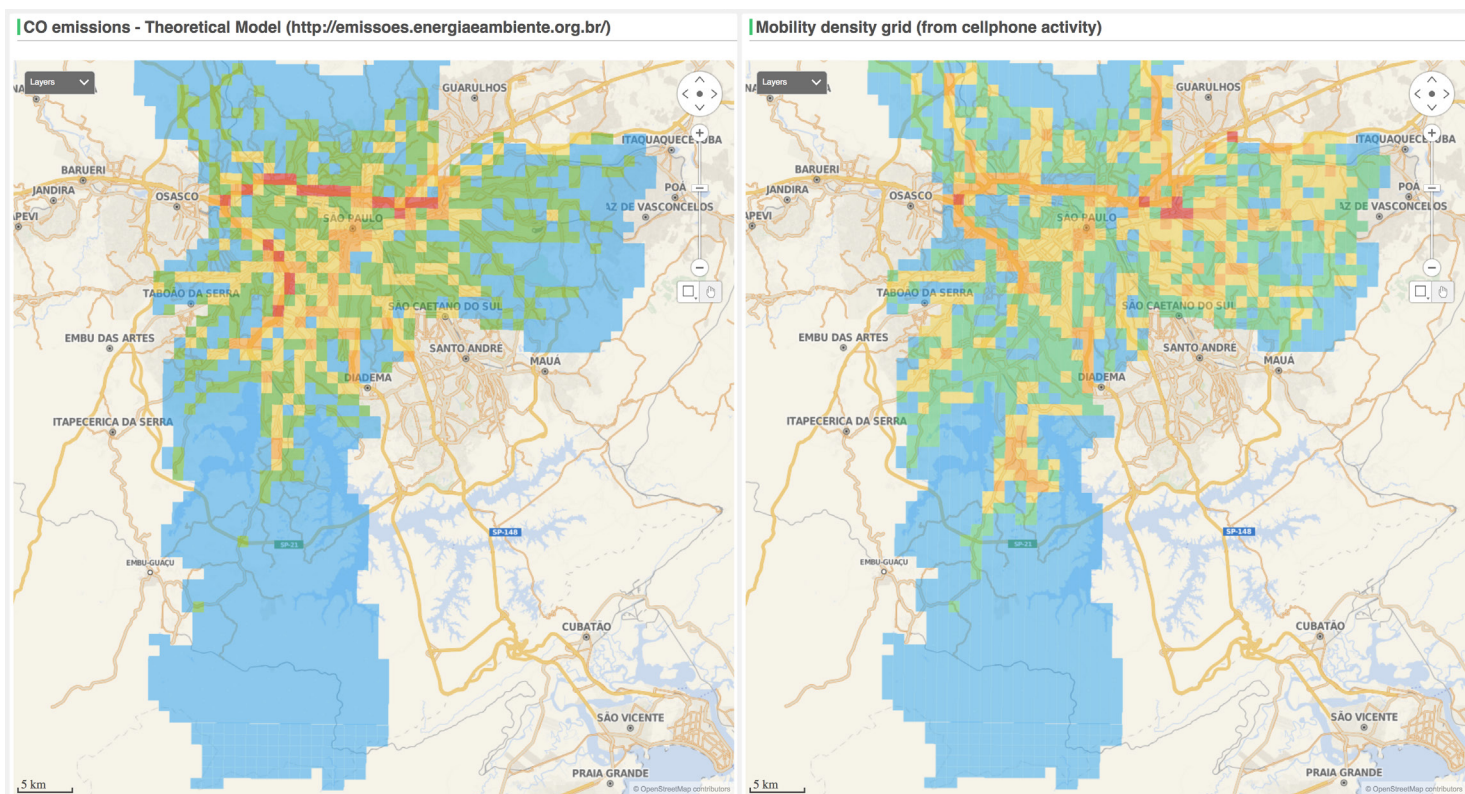
No entanto, o gerenciamento da qualidade do ar pode custar caro e a análise do trânsito pode ser um processo relativamente manual, utilizando dados de entrevistas feitas nas rodovias ou contadores manuais. A infraestrutura de monitoramento da qualidade do ar é limitada, na maioria das cidades, devido ao alto custo de instalação e operação de estações de medição fixas. Isso significa que os dados são insuficientes, tornando difícil para os órgãos públicos a coleta de dados precisos, os quais são necessários para identificar onde estão os principais pontos críticos de poluição do ar.

Em São Paulo, o governo local se baseia em um modelo teórico para emissões de CO₂ com base em medições limitadas de trânsito em horários fixos do dia. Os dados móveis dão muito boa concordância a este modelo e podem ser generalizados e refinados para fornecer previsões mais úteis (veja a Figura 1)

2. <http://www.who.int/airpollution/en/>

3. Miraglia S., Gouveia N. Costs of air pollution in Brazilian metropolitan regions. Cienc. Saude Colet. 2014;19:4141-4147

Figura 1 O painel da Telefonica mostra o forte acordo entre as medições de trânsito feitas a partir de dados móveis para as previsões teóricas de CO₂ do governo



Como os dados móveis podem ajudar

A Telefônica Brasil está colaborando com a cidade de São Paulo para fazer uma correlação entre as informações obtidas de sensores de qualidade do ar, sensores de trânsito e estações meteorológicas com os dados gerados por redes móveis. Essa abordagem permite chegar a conclusões relevantes de forma mais econômica quando comparado com a observação direta de níveis de poluição, e de forma mais dinâmica e constante do que seria possível por meio de coleta de dados mais tradicionais.

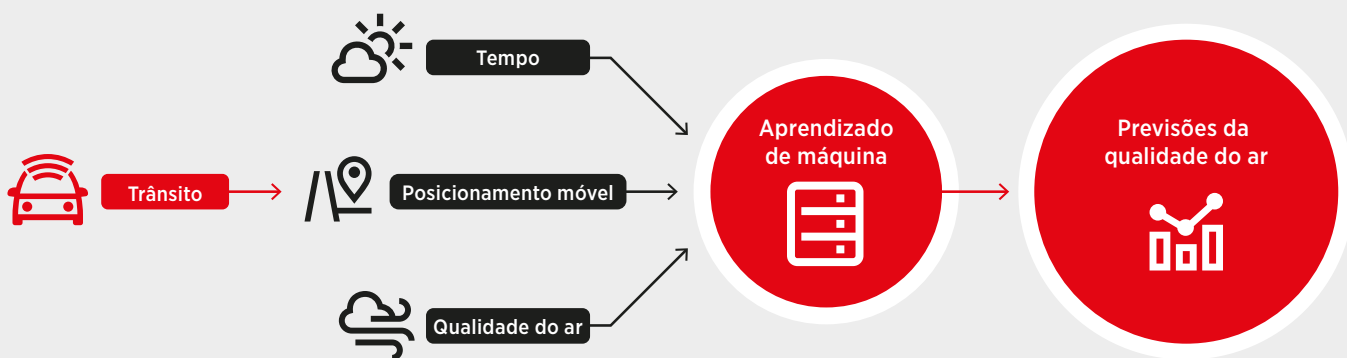
Para compensar o número limitado de sensores na cidade, a Telefônica aproveita os dados móveis para servir como proxy para observações sobre o volume e o fluxo de veículos. Os cientistas de dados da Telefônica LUCA criaram um conjunto de algoritmos para calcular estimativas de mobilidade e trânsito e um painel que exibe os níveis de poluição em toda a cidade. Os algoritmos utilizam machine learning e dados de clima, trânsito e poluição para extrapolar o mapeamento sobre a cidade inteira (ver Figura 2).

Os dados da Telefônica vêm da atividade da rede, como as chamadas de voz, SMS ou acesso à Internet. Esses dados podem ser agregados e incrementados para criar um índice que represente toda a população da cidade. Segundo a Telefônica, tal índice geralmente tem uma correlação (de até 94%) com o trânsito realmente observado. Para preservar a privacidade do indivíduo, o anonimato é feito através da remoção do MSISDN, da adição de ruído, agregação e extrapolação da fatia de mercado da Telefônica para toda a população e pela normalização⁴.

A Telefônica também está testando a solução de dados em Madri para criar uma abordagem robusta que possa ser utilizada para diferentes lugares.

4. Por exemplo, fornecendo desvios da mobilidade média em cada localidade – a Telefônica calcula o índice de mobilidade de cada uma das estações de medição. A correlação varia de 89% a 97% dependendo da estação”. 94% é a média entre as diferentes correlações.

Figura 2 Combinando os dados de poluição, de trânsito e de telecomunicações para mapear e prever a qualidade do ar



Dados de posicionamento móvel são usados como proxy para dados de trânsito, permitindo que o trânsito seja medido em toda a cidade, em vez de nos locais da rede limitada de sensores de trânsito. Os dados móveis, juntamente com dados meteorológicos e históricos de qualidade do ar, são combinados em um modelo de aprendizado de máquina para gerar previsões da qualidade do ar.

O impacto da solução de dados móveis

A implementação em São Paulo mostrou que big data baseado em dados móveis é uma ferramenta poderosa para prever e analisar padrões de trânsito e a poluição do ar resultante dentro de uma cidade. No geral, os dados móveis permitem um mapeamento mais granular de poluição da cidade do que poderia ser realizado por meio de uma rede de sensores estáticos caros.

Essa prova de conceito em São Paulo destaca como os dados de rede móvel, correlacionados com dados de qualidade do ar, de clima e de outras fontes, podem

prever, com antecedência, níveis de poluição em 24 a 48 horas. Assim, as autoridades locais podem tomar medidas preventivas, caso as emissões de NO₂ comprometam a saúde humana. Tal estratégia pode reduzir significativamente os custos, aumentar a precisão e fornecer informações importantes aos responsáveis sobre como definir as políticas de transporte urbano ou avaliar os níveis de poluição. Além disso, terceiros podem utilizar uma API para conseguir acesso direto a dados agregados e anônimos de telecomunicações e aos insights sobre mobilidade para fins próprios.



O desenvolvimento de uma proposta sustentável

Uma lição fundamental que resulta de tudo isso é a seguinte: para criar uma implementação sustentável, os governos locais e as operadoras de telefonia móvel tem que colaborar para combinar seus conjuntos de dados e para analisar novos modelos que possam gerar melhores previsões, e também para definir as políticas a serem desenvolvidas ou informadas com base nesses insights.

Para tirar vantagem de todo o potencial de big data baseado em dados móveis, é preciso que as

administrações públicas modifiquem o orçamento anual de projetos para incluir o gerenciamento desses serviços e para investir nas competências relevantes. Isso contribuirá para diminuir os custos associados com a saúde, o desenvolvimento da infraestrutura local, e a mobilidade urbana. Algumas administrações já reservaram orçamento para a análise de dados e estudos de casos, incluindo especificamente conjuntos de dados móveis para melhorar seus modelos⁵.

5. Plano de Inovação para Transportes e Infraestruturas 2017 – Ministério de Obras Públicas e Transportes – Governo da Espanha



Os próximos passos

A Telefônica identificou uma série de próximos passos possíveis:



Aprimorar as ferramentas em termos de robustez e relevância, como por exemplo: conseguir fontes de dados adicionais dentro dos processos operacionais das autoridades locais e melhorar a qualidade dos modelos preditivos.



Ajudar os municípios a continuarem com a otimização da rede de estações que monitoram os níveis de poluição e de trânsito nas cidades e mais extensivamente no país para melhorar os algoritmos e as capacidades de previsão.



Desenvolver um roteiro para transformar a solução de São Paulo em produtos mais escaláveis que poderiam ser implementados em outras regiões geográficas. A Telefônica está desenvolvendo uma estratégia parecida para cobrir o município de Madri, capital da Espanha.



Desenvolver melhor as fontes de renda provenientes de terceiros para acessar os dados agregados de telecomunicações através de uma API para outros fins que requerem ou abrangem a mobilidade e/ou insights ambientais para criar valor tanto econômico (por exemplo, campanhas de marketing direcionadas) como social (por exemplo, preparação para desastres).

Assista aos nossos vídeos, conheça mais sobre a nossa iniciativa e entre em contato para mais informações: bd4sg@gsma.com

gsma.com/betterfuture/bd4sg



@GSMA

#BetterFuture #BD4SG