

O fosso digital entre a Amazônia e demais regiões do Brasil: Análise baseada em redes bayesianas e Sistemas de Informação Geográfica

The digital divide between the Amazon and other regions of Brazil: Analysis based on Bayesian networks and Geographic Information Systems

DOI:10.34117/bjdv7n9-485

Recebimento dos originais: 07/08/2021

Aceitação para publicação: 27/09/2021

Francinete Furtado da Cunha

Graduada pela Universidade Federal Rural da Amazônia
Mestranda em Ciência da Computação na Universidade Federal do ABC
Avenida dos Estados, 5001 Bairro Santa Terezinha
Santo André – SP, Cep: 09.210-580
E-mail: francinete.furtado@ufabc.edu.br

Barbara Kathellen Andrade Porfirio

Mestre em Ciência da Computação na Universidade Federal do ABC
Avenida dos Estados, 5001 Bairro Santa Terezinha
Santo André – SP, Cep: 09.210-580
E-mail: barbara.andrade1901@gmail.com

Thaynara Cristina Maia dos Santos

Graduada pela Universidade Federal Rural da Amazônia
Mestranda em Ciência da Computação na Universidade Federal do ABC
Avenida dos Estados, 5001 Bairro Santa Terezinha
Santo André – SP, Cep: 09.210-580
E-mail: thaynara.maia10@gmail.com

Thalynny Rayssy Rabelo da Silva

Graduada pela Universidade Federal Rural da Amazônia
Engenheira de Software na PagSeguro
Avenida Brigadeiro Faria Lima, N° 1384 Bairro Jardim Paulistano
São Paulo-SP, Cep: 01.452-002
E-mail: thalynny.r@gmail.com

Aleksandra do Socorro da Silva

Doutorado pela Instituição Universidade Federal do Pará
Universidade Federal Rural da Amazônia
Avenida Presidente Tancredo Neves, N° 2501 Bairro Terra Firme
Belém-Pará, Cep: 66.077-830
E-mail: aleksandra.silva@ufra.edu.br

Silvana Rossy de Brito

Doutorado pela Instituição Universidade Federal do Pará
Universidade Federal Rural da Amazônia
Avenida Presidente Tancredo Neves, N° 2501 Bairro Terra Firme

Belém-Pará, Cep: 66.077-830
E-mail: silvana.rossy@ufra.edu.br

RESUMO

Atualmente o grande desafio é a ampliação do acesso à Internet para aqueles que ainda estão excluídos digitalmente, sendo necessário um esforço em conjunto de todos os governantes, para que se avance no desenvolvimento de uma sociedade mais incluída digitalmente. Neste intuito, é realizado uma análise quantitativa e qualitativa da posse de microcomputador com e sem Internet nos municípios da Amazônia Legal Brasileira, aplicando uma metodologia já consolidada para medir a exclusão digital nesses municípios. Utilizou-se de dados obtidos do Censo Demográfico do IBGE, no ano de 2010 e Atlas de Desenvolvimento Humano do Brasil do ano de 2013, tendo como referência o ano de 2010, com técnicas de redes bayesianas e Sistemas de Informação Geográfica. Os resultados demonstram como os indicadores de educação e renda continuam sendo determinantes para o baixo índice de posse de microcomputador com e sem internet na Amazônia.

Palavras-chave: redes bayesianas, dados espaciais, microcomputador, acesso à internet, sistemas de informação geográfica.

ABSTRACT

Currently, the great challenge is the expansion of Internet access for those who are still digitally excluded, requiring a joint effort by all government officials to move forward in the development of a more digitally included society. For this purpose, a quantitative and qualitative analysis of the possession of microcomputers with and without Internet is carried out in the municipalities of the Brazilian Legal Amazon, applying a consolidated methodology to measure the digital exclusion in these municipalities. We used data obtained from the IBGE Demographic Census in 2010 and the Human Development Atlas of Brazil in 2013, with the year 2010 as a reference, with Bayesian networks techniques and Geographic Information Systems. The results demonstrate the enormous distances between the indicators of the Amazon region and the other municipalities of Brazil, as the education and income indicators continue to be determinant for the low index of possession of microcomputer with and without internet in the Amazon.

Keywords: Bayesian networks, spatial data, microcomputer, internet access, geographic information systems.

1 INTRODUÇÃO

Dado um ambiente social em que não existam disparidades socioeconômicas, o uso de Tecnologias de Informação e Comunicação (TICs) parece ser promissor e possuir um potencial fantástico. Para muitos países em desenvolvimento, como o Brasil, a exclusão digital é um limitador de acesso aos serviços de educação, saúde, governança pública. Compreender as disparidades regionais é importante para fundamentar a construção de políticas públicas para que a exclusão não seja perpetuada, criando-se um

abismo ainda maior entre os que têm e os que não têm acesso à informação e aos serviços de governo e da sociedade.

Em 2012, de acordo com Neri (2012), o principal motivo da exclusão digital era a falta de conhecimento (33%), sendo que a grande maioria com acesso está localizada nas capitais mais desenvolvidas em IDHM (Índice de Desenvolvimento Humano por Município) e o segundo é a falta de habilidade de usar a Internet (31%), onde se encontram alguns dos municípios com índice de desenvolvimento econômico e social mais baixo do país, ambos decorrentes dos problemas educacionais vigentes no país (NERI, 2012). Para quantificar essa desigualdade, os autores usaram métodos quantitativos, expressos através de indicadores e estatísticas, abordagem comumente aplicada em estudos, como o Mapa de Inclusão Digital.

De acordo com a União Internacional de Telecomunicações (ITU, do inglês International Telecommunication Union) (ITU, 2019), 49% da população mundial ainda não possui acesso à Internet e a maior parte dessas pessoas vivem em países em desenvolvimento. Essa é a situação de muitas comunidades ribeirinhas, quilombolas e indígenas, que vivem às margens dos rios e em áreas rurais isoladas da Amazônia legal brasileira. Sob o ponto de vista das definições de direitos fundamentais, o acesso à internet deve ser classificado como um direito fundamental, com o objetivo de gerar iniciativas do Estado que aumentem o acesso à rede e, ao mesmo tempo protejam os direitos dos que a acessam (SANTO E POMIN, 2021). Para os autores, há dois mundos completamente diferentes: um composto por pessoas que tem acesso à rede e outro, por indivíduos que não tem; e o Estado não pode ficar inerte se esse é um direito humano fundamental – que garante a efetivação de outros direitos – devendo assumir o importante papel de propor e conduzir políticas públicas para reduzir a exclusão digital no país.

Este trabalho apresenta uma abordagem para avaliar a centralização de acesso às TICs, efetuadas a partir dos municípios, buscando dimensionar o tamanho dessa desigualdade no Brasil, sob a ótica dos amazônidas. Considerando a importância do acesso às TICs e a carência de estudos que qualifiquem o fosso digital nos municípios da Amazônia em relação aos municípios das demais regiões do Brasil, este estudo tem o objetivo analisar de forma quantitativa e qualitativa a associação entre o acesso às TICs e variáveis representativas das dimensões de educação, renda, longevidade e Produto Interno Bruto (PIB) nesses municípios.

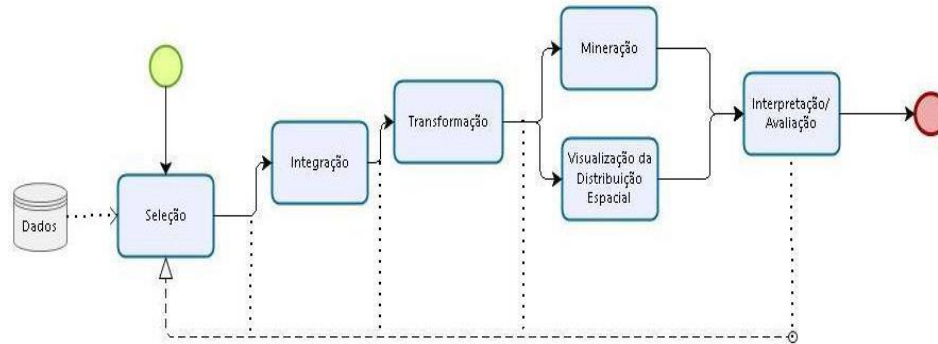
Entre os métodos estatísticos, destaca-se o aprendizado Bayesiano, que utiliza um modelo probabilístico baseado no conhecimento prévio do problema, o qual é combinado

com exemplos de treinamento para determinar a probabilidade final de uma hipótese (REZENDE, 2003; ROCHA, 1999). O estudo da associação, utilizando redes Bayesianas permite encontrar associações entre fenômenos estudados, mas não considera as distâncias geográficas entre os locais onde esses fenômenos ocorrem. De fato, a proximidade das localidades pode influenciar para mais ou para menos, a ocorrência do fenômeno da exclusão digital. Com o objetivo de identificar esse aspecto, este trabalho faz uso do Sistema de Informação Geográfica (SIG) para agregar a análise de dados espaciais às considerações analíticas alcançadas com a mineração com redes bayesianas. Assim, aplicamos tecnologias de mineração de dados de forma conjunta com a análise da distribuição espacial utilizadas em trabalhos prévios (BRITO, 2016; BRITO et al., 2016; SILVA et al., 2016; SILVA, 2016), que combina a estratégia de busca de associação entre variáveis e o uso de SIG. Como resultado, apresentamos as associações que influenciam na posse de microcomputador com e sem internet e as diferenças probabilísticas entre a região amazônica e os demais estados do Brasil. Para a discussão, apresentamos dados atualizados que corroboram as análises realizadas.

2 MATERIAIS E MÉTODOS

Nesta pesquisa aplicada, realizamos um estudo descritivo, transversal, de caráter quantitativo, com foco na exposição de características de determinada população e fenômenos, através da identificação de associações entre importantes indicadores como forma de apoiar análises qualitativas por parte dos especialistas. A metodologia é estendida a partir de uma adaptação do processo de Descoberta de Conhecimento em Banco de Dados (do inglês Knowledge Discovery in Databases) que é um processo de várias etapas, não trivial, interativo e iterativo, para identificação de padrões compreensíveis, válidos, novos e potencialmente úteis a partir de grande conjunto de dados (FAYYAD et al., 1996). Nossa abordagem inclui a análise da distribuição espacial, identificando, no estudo, se os municípios pertencem à região Amazônica ou não. Os procedimentos adotados são apresentados, de forma integrada, na Figura 1.

Figura 1. Metodologia adotada para análise.



Foram utilizadas as seguintes etapas (Figura 1): (1) seleção das variáveis do domínio do problema e relativas aos municípios da Amazônia e de outras regiões do país, a partir de diferentes fontes de dados; (2) integração dos dados em uma única base; (3) tratamento (transformação) dos dados para permitir a execução de algoritmos de mineração de dados e de visualização da distribuição espacial; (4) Mineração e visualização da distribuição espacial dos dados; e, (5) interpretação/avaliação dos dados, confrontando as informações a partir da visualização dos dados com as regras de associação encontradas.

Na primeira etapa de pré-processamento é realizada a seleção dos dados. Como fontes de dados, utilizou-se: (1) os microdados da base da Pesquisa de Domicílios do Censo Demográfico do IBGE tendo como referência os municípios brasileiros e o ano de 2010, por este ser o censo mais recente que inseriu a pergunta sobre “O acesso à microcomputador (*mc*) e acesso à microcomputador com Internet (*mc_int*)” nos domicílios; (2) os componentes do *Índice de Desenvolvimento Humano Municipal (idhm)*, para o ano de 2010, referentes à *educação (idhm_e)*, *renda (idhm_r)*, *longevidade (idhm_l)* e *pib*, obtidos a partir da plataforma de consulta Atlas do Desenvolvimento Humano no Brasil (2013). Na base do Censo Demográfico, a amostra abrangeu 57.320.474 domicílios em 5.565 municípios, sendo 771 localizados na Amazônia.

A segunda etapa, de integração dos dados, é necessária porque os dados provêm de diferentes fontes. Uma vez que os dados possuem um identificador em comum, que é o código do município fornecido pelo IBGE, eles são integrados em uma única base de dados.

Na terceira etapa, de transformação, os dados foram tratados para que possam servir de entrada para os algoritmos de mineração e de distribuição espacial. Dessa forma, as variáveis *mc*, *mc_int*, *idhm_e*, *idhm_r*, *idhm_l*, *pib* por serem variáveis contínuas,

passaram pelo processo de discretização por frequência, que permite com que fossem definidos intervalos de mesma frequência para a distribuição dos valores de cada variável em faixas de valores. A variável local é uma variável discreta, onde só é possível a existência de dois estados: ou o município pertence à Amazônia ("Amazônia") ou não pertence ("Outra região").

Na quarta etapa é aplicado um algoritmo de mineração de dados, com o objetivo de medir a associação entre as variáveis estudadas, e de visualização da distribuição espacial do fenômeno de posse de microcomputador e microcomputador com internet. Para a mineração, foi utilizada a técnica de redes bayesianas. Segundo Chen (2001) rede Bayesiana “é um grafo acíclico dirigido consistindo de nós, que representam variáveis aleatórias, e setas, que representam relações probabilísticas entre essas variáveis”. Para o aprendizado da estrutura da rede bayesiana foi aplicado o algoritmo de busca heurística K2 (Cooper e Herskovits 1992), tendo como entrada os dados referentes às variáveis do domínio do problema (*mc*, *mc_int*, *idhm_e*, *idhm_r*, *idhm_l*, *pib*, *local*). O K2 aplica um método de pontuação bayesiano para cada estrutura testada e seu objetivo é encontrar a estrutura de rede bayesiana mais provável no espaço de busca. A escolha para o K2 foi devido ao seu desempenho em termos de complexidade computacional e resultados precisos, obtidos quando uma ordenação de variáveis adequada é fornecida. Nesse estudo, a ordenação exigiu que a variável alvo (*mc_int*) fosse mantida após todas as demais variáveis.

Uma vez gerada a rede bayesiana, foi utilizada a inferência estatística a fim de estimar a posterior distribuição dos parâmetros. A tarefa básica da inferência consistiu em computar a distribuição da probabilidade condicional utilizando as informações quantitativas da rede bayesiana, ou seja, para um conjunto de variáveis de consulta (*Consulta*), dado os valores de um evento observado (*Evidências*), o sistema computa $P(Consulta|Evidências)$. Para avaliar a relevância das regras de associação produzidas, foi utilizado o modelo suporte-confiança (AGRAWAL et al. 1993). Assim, foram selecionadas as regras que apresentaram um suporte mínimo ($Sup \geq 5\%$) e uma confiança mínima ($Conf \geq 65\%$).

Em paralelo à aplicação de redes bayesianas, ainda na quarta etapa, foram gerados diferentes mapas para promover a visualização da distribuição espacial dos percentuais – essa abordagem permite apresentar de forma visual o resultado da associação entre as variáveis do domínio do problema, de forma geográfica no mapa temático. Para obter essa visualização foi aplicado uma base de dados através do software QGIS (2021).

A última etapa envolveu a avaliação e interpretação desses resultados a fim de extrair conhecimentos, para possíveis tomadas de decisão. A estratégia envolveu analisar os dados focando principalmente nas comparações de municípios da Amazônia com outras regiões do país.

3 RESULTADOS E DISCUSSÃO

As regiões Norte e Nordeste são as que apresentam a segunda maior diferença em pontos percentuais em relação ao Brasil, conforme dados do IBGE, tanto em relação à posse de microcomputador quanto a posse de *microcomputador com internet*, conforme já apontado por vários autores (NERI, 2012; BRITO, 2016; BRITO et al, 2016, SILVA et al 2016).

Tabela 1 – Posse de microcomputador e microcomputador com Internet em municípios do Brasil, Regiões e Amazônia.

Média	Posse de microcomputador (%)	Diferença Percentual (%)	Posse de <i>microcomputador com internet</i> (%)	Diferença Percentual (%)
Brasil	38,31	-	30,73	-
Norte	22,69	-15,62	15,44	-15,29
Nordeste	21,22	-17,09	16,78	-13,95
Sudeste	47,99	9,68	39,59	8,86
Sul	46,10	7,79	35,52	4,79
Centro-Oeste	39,18	0,87	31,44	0,71
Amazônia	17,8	-20,51	17,5	-13,23

A Região Norte, por exemplo, apresenta uma diferença de -15,62% e -15,29%, dados de IBGE (2010), para a posse de microcomputador com e sem Internet, respectivamente, em relação ao Brasil. Na Tabela 1, além das médias do Brasil e regiões, há também a média da Amazônia, que apresenta a maior diferença percentual em relação às demais regiões do país: Sul e Sudeste estão com 7,79% e 9,68% nesta ordem, acima da média Brasileira na posse de microcomputador e 4,79% e 8,86% sobre a posse de microcomputador com internet, respectivamente. Sobre a região Centro-Oeste, que possui uma pequena quantidade de municípios que fazem parte da Amazônia legal brasileira, ela apresenta índices acima da média Brasileira.

Na sequência, as variáveis contínuas foram discretizadas para que a base de dados fosse submetida ao algoritmo K2. A discretização foi realizada em quatro faixas de valores para cada variável, denominando cada indicador como: muito baixo, baixo, médio e alto para cada faixa de valor, observados de cima para baixo, respectivamente. Na

Tabela 2, as variáveis *pib* e *idhm_l* não são apresentadas porque não apresentaram associação direta com as variáveis alvo do problema.

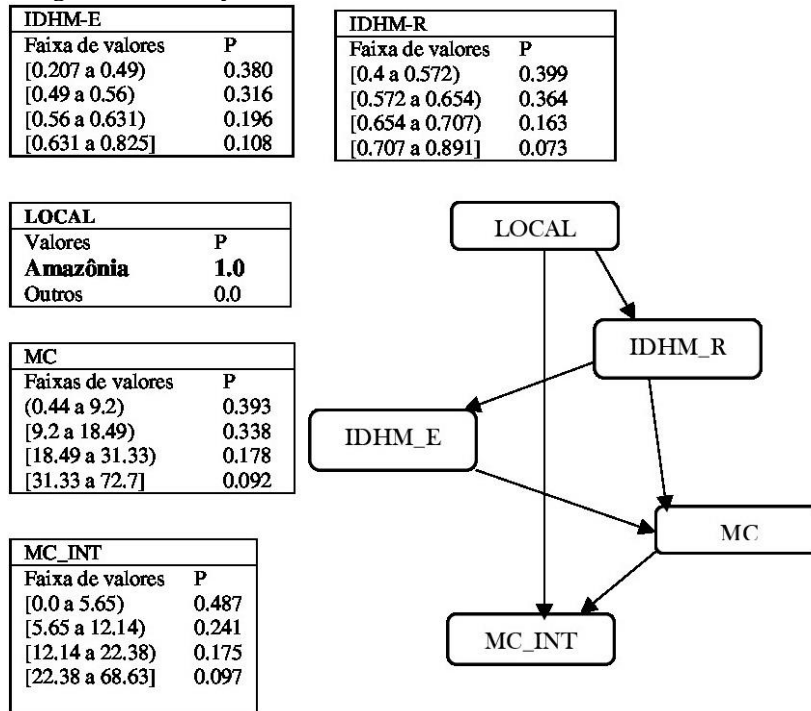
Tabela 2. Discretização por frequência das variáveis contínuas do domínio do problema -

ESCALA	<i>idhm_r</i>	<i>idhm_e</i>	<i>mc</i> (%)	<i>mc_internet</i> (%)
Muito baixo	[0,400 - 0,572)	[0,207 - 0,490)	[0,44 - 9,2)	[0 - 5,65)
Baixo	[0,572 - 0,654)	[0,490 - 0,560)	[9,2 - 18,49)	[5,65 - 12,14)
Médio	[0,654 - 0,707)	[0,560 - 0,631)	[18,49 - 31,33)	[12,14 - 22,38)
Alto	[0,707 - 0,891]	[0,631 - 0,825]	[31,33 - 72,7]	[22,38 - 68,63]

É possível observar as tabelas de probabilidade dos nós da rede gerada, colocando-se em evidência Local = "Amazônia" (Figura 2). Assim, em cada tabela cuja variável tem associação com local, haverá a propagação das probabilidades condicionais, permitindo-se realizar inferências relevantes para a compreensão do fenômeno da posse de microcomputador e microcomputador com Internet na Amazônia e nas demais regiões do país.

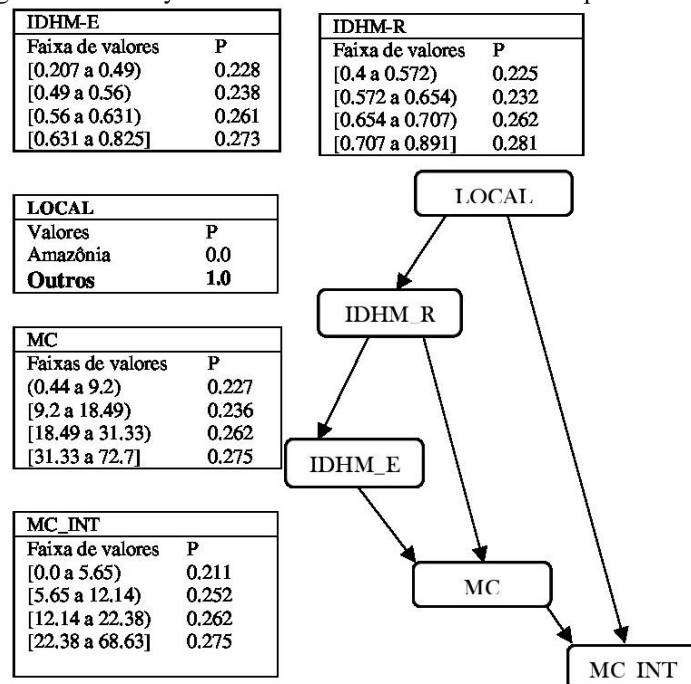
De fato, dada a evidência de que um município está situado na região da Amazônia (Figura 2), a probabilidade de o município apresentar as variáveis *idhm_r* e *idhm_e* na faixa "Muito baixo" é de 39,9% e 38,0%, respectivamente; por outro lado, quando se trata de municípios em outras regiões do país (Figura 3) essa probabilidade cai para 22,5% e 22,8%, respectivamente. Nesse cenário, a probabilidade das variáveis *mc* e *mc_int* estarem faixa mais elevada ("Alta"), quando se trata de um município da Amazônia, é de 9,2% e 9,7% nesta ordem e quando se trata de municípios de outras regiões (Figura 3), essa probabilidade aumenta para 27,5%, nos dois casos.

Figura 2. Rede bayesiana – em evidência a variável Local = “Amazônia”



Quando a evidência é a de que o Local são “outros municípios do Brasil”, a probabilidade de estar na faixa mais baixa de renda (< 0,57%) e educação (< 0,49%) é de 22,5% e 22,8%, respectivamente; a probabilidade de posse de microcomputador desses municípios cair na faixa mais alta de valores de valores (> 31,33%) é de 27,5%.

Figura 3. Rede bayesiana – em evidência “Outros municípios do Brasil”

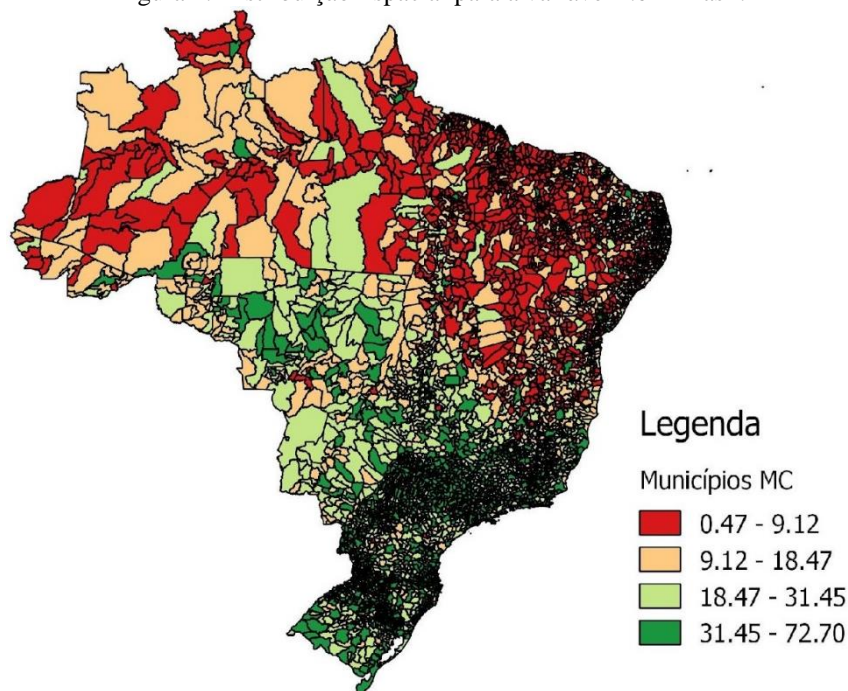


Quando colocamos a evidência outros municípios do país, para a faixa mais elevada de renda e educação, a probabilidade de *IDHM_R* ($> 0.707\%$) e *IDHM_E* (> 0.631), chega a 0,281% e 27,3%, respectivamente. Com esse resultado, a probabilidade de acesso a posse de *mc_int* desses municípios na faixa mais elevada de valores ($> 22,38\%$) é de 27,5%.

A partir da análise dos dados, foram gerados dois mapas pelo software QGIS, para visualizar de forma percentual, a atuação nos municípios brasileiros no acesso às TICs, especificamente em relação à posse de microcomputador e microcomputador com internet, sendo possível obter informação espacial dessa ocorrência, de acordo com a faixa de valores estabelecidos no processo de discretização.

Os indicadores de posse de microcomputador e microcomputador com Internet, quando apresentados na forma de distribuição espacial, demonstram claramente que as regiões Norte e Nordeste concentram os municípios com os percentuais mais baixos da posse desses recursos (Figura 4 e Figura 5).

Figura 4. Distribuição Espacial para a variável *mc* – Brasil.

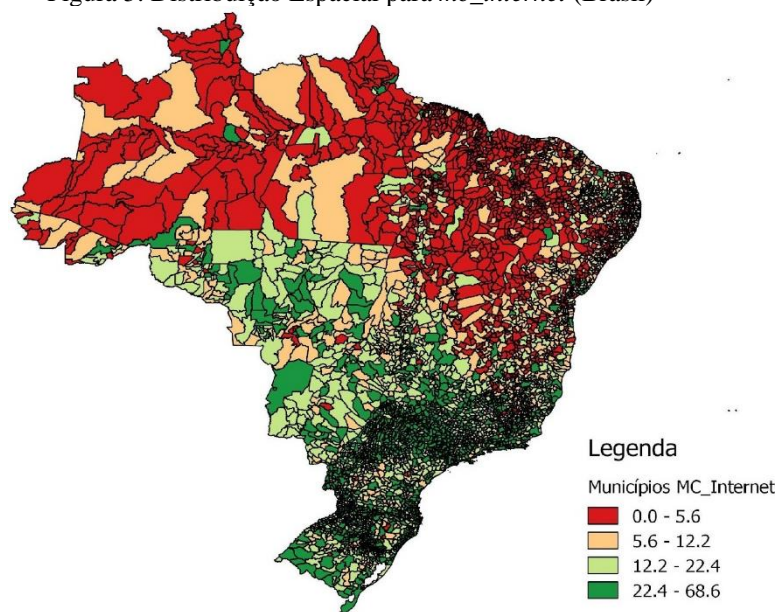


Fonte: Elaborado pelas Autoras.

As fontes de dados contaram com uma limitação importante: os dados, por município somente estão disponíveis no Censo Demográfico de 2010. Outra pesquisa que contemple as variáveis de posse de microcomputador e microcomputador com acesso à

Internet é a Pesquisa Nacional por Amostra de Domicílios (PNAD). Entretanto, a PNAD é uma pesquisa amostral, realizada em todas as regiões do Brasil, mas não permite obter dados agregados por municípios e não tem valor absoluto sobre toda a população, como o Censo Demográfico. Para apoiar a discussão, apresentamos a seguir, informações mais atuais que corroboram nossas conclusões.

Figura 5. Distribuição Espacial para *mc_internet* (Brasil)



Fonte: Elaborado pelas Autoras.

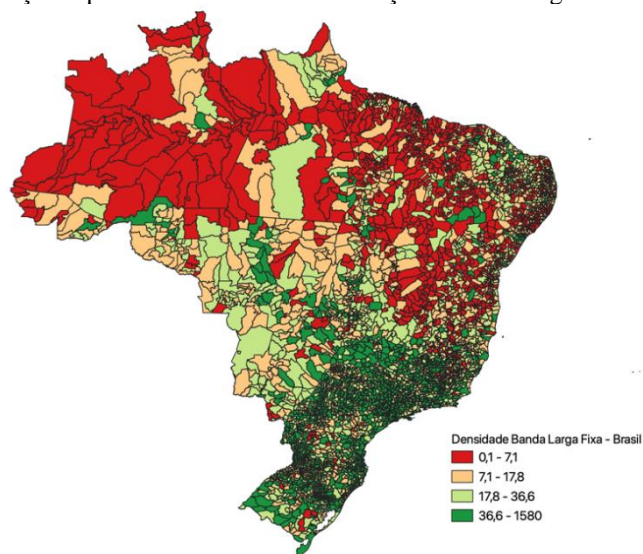
Para apresentar dados mais atuais, utilizamos informações da Agência Nacional de Telecomunicações (Anatel). Informações da ANATEL permitem mensurar a disponibilidade e a qualidade do acesso à Internet nos municípios brasileiros, ainda que seja uma tarefa complexa que envolve fatores diretos e indiretos de prestação dos serviços de telecomunicações. A ANATEL informa mensalmente a quantidade de acessos existentes no Brasil relativos aos principais serviços de telecomunicações. Por exemplo, o indicador densidade representa os acessos dos serviços de banda larga fixa e do serviço móvel pessoal. A densidade é um indicador bem estruturado porque é calculado a partir da população do município. Além desse indicador, há o indicador de computador com acesso à Internet, provido pela PNAD contínua e que apresenta o percentual, por unidade federativa, anualmente, da posse de computador com acesso à Internet.

As tecnologias envolvidas no acesso à Internet podem, em termos gerais, ser divididas em: (1) baseadas em infraestrutura fixa, e (2) e baseadas em infraestrutura móvel. Para nossas considerações analíticas consideramos os dados de duas categorias de

tecnologias para construir mapas de distribuição geográfica da densidade, tendo como referência o mês de junho de 2020 para todos os municípios brasileiros.

Por exemplo, quando observamos a distribuição espacial da densidade da banda larga fixa no Brasil, fica claro que a tecnologia permanece concentrada nas regiões Sul, Sudeste e Centro-Oeste (Figura 6). Os valores desse indicador, que são contínuos, foram discretizados em 4 faixas de valores, de forma que cada faixa representa 25% dos municípios brasileiros – mesma técnica utilizada para todo o processo de discretização dos dados minerados.

Figura 6. Distribuição espacial da densidade do serviço de banda larga fixa – Brasil



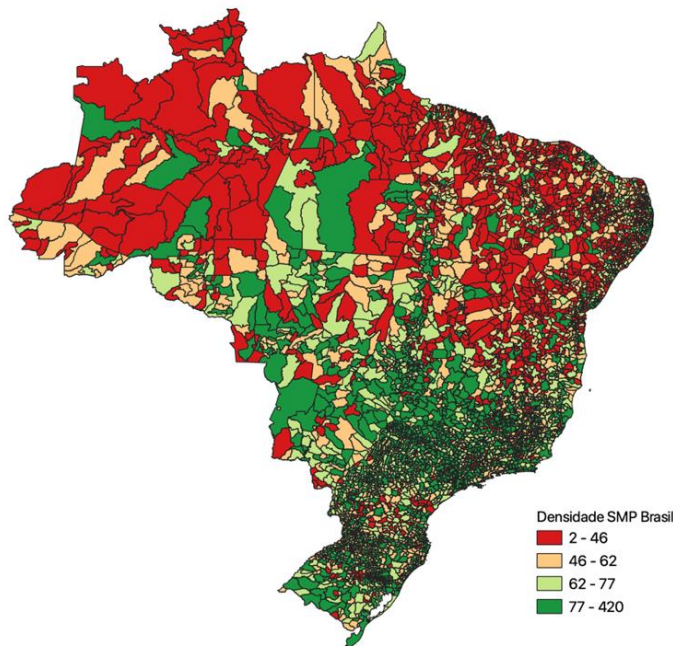
Fonte: Elaborado pelas Autoras.

Como resultado (Figura 6), a menor faixa da densidade de banda larga fixa (0,1 a 7,1) está principalmente concentrada na região Norte do país – cenário que parece não ter sido alterado, quando observado em comparação com as demais regiões do país. Enquanto esse indicador, para o Brasil, é de 48,8, a região Norte mantém o valor mais baixo (25,4) dentre as demais regiões e o Pará, da região Norte, é o estado com a menor densidade em banda larga fixa do Brasil, registrando o valor de 18,5.

Na última década, as conexões móveis, principalmente o 3G e 4G, que utilizam a infraestrutura das linhas telefônicas móveis, avançam sobrepondo a expansão da banda larga em todo o país. O Serviço Móvel Pessoal (SMP) requer menos gastos com infraestrutura se comparado à redes fixas, facilitando tanto o acesso em regiões sem infraestrutura básica de telecomunicações, principalmente em regiões rurais ou ribeirinhas, embora, por depender do tamanho do espectro da radiofrequência e sendo este limitado, sua expansão pode sobrecarregar a capacidade da rede, afetando a qualidade

do serviço. No Brasil, a densidade do SMP é de 95,7, sendo que as regiões Centro-Oeste e Sudeste concentram a maior densidade neste serviço (104,1 e 103,8 respectivamente). As regiões Norte e Nordeste são as com a menor densidade (84,6 e 83,9), seguidas pela região Sul (96,6). O mapa da Figura 7 apresenta a distribuição espacial desse indicador nos municípios Brasileiros. O Pará é o Estado da região Norte com a menor densidade no SM (80).

Figura 7. Distribuição espacial da densidade do SMP – Brasil



Fonte: Elaborado pelas Autoras.

Dessa forma, podemos concluir que mesmo observando dados mais atuais do acesso e densidade de acesso à internet no Brasil, o cenário que identificamos quando usamos a base do Censo Demográfico de 2010 parece não ter mudado quando se trata de analisar as desigualdades, ou o fosso digital, no país. É importante ressaltar a importância da realização do Censo Demográfico para revelar avanços e novos desafios para a região, pois somente através do Censo é possível avaliar as condições de acesso nos domicílios, por município.

4 CONCLUSÕES

Este estudo ressalta a associação entre a posse às TICs e variáveis representativas das dimensões de *renda*, *educação*, *PIB* e *longevidade*, esta que influencia indiretamente no acesso nos municípios brasileiros. Os percentuais de municípios que tem acesso às TICs apresentaram-se maiores em outros municípios do Brasil, enquanto na Amazônia,

esse indicador permanece muito reduzido, com grande dependência nas variáveis de *IDHM_E* e *IDHM_R*. Além disso, podemos perceber que o valor do *PIB* do município não demonstrou ser uma variável com associação significativa para o estudo. Apesar de muitos dos municípios da Amazônia apresentarem *PIB* alto, a posse de microcomputador e microcomputador com Internet foi considerada baixa nesses municípios, o que pode significar a necessidade de investimentos em recursos na estrutura desses municípios – o que parece ser confirmado mesmo com dados atuais da densidade demográfica.

De modo geral, os possíveis fatores que têm um alto grau de influência no baixo acesso à Internet nos municípios da Amazônia são permanecem: municípios localizados na zona rural, com baixa renda e níveis educacionais, com fraco suporte à infraestrutura de telecomunicações.

É importante ressaltar a importância dos estudos que avaliam a capacidade dos domicílios em suportar o uso de plataformas digitais na atualidade e que demonstrem as desigualdades enfrentadas. De acordo com Santana et al. (2020), as instituições de ensino tiveram que evoluir rapidamente suas ferramentas e plataformas digitais para garantir uma entrega educacional ininterrupta aos estudantes isolados, pois a situação de pandemia pelo Covid-19 forçou as instituições ao uso de ferramentas *on-line*. Da mesma forma, exigiu dos estudantes equipamentos e acesso à Internet apropriados para acessar tais conteúdos e interações.

O desafio não é apenas para a disseminação do conteúdo digital formal provido pelas escolas. Relatam Godin et al. (2020) que na comunidade de Turé, uma reserva extrativista Terra Grande Pracuúba localizada no alto Rio Piriá, município de Currealinho Marajó, no Estado do Pará, Amazônia brasileira, muitas adolescentes vivenciam problemas levando a evasão escolar, pois não dispõe de uma estrutura familiar que possa contribuir de forma positiva nessa fase, vivendo em situação de extrema pobreza. Para os autores, no Estado do Pará, em muitos municípios da região não há condições favoráveis para a vida da sua população: não há saneamento, rede de saúde pública e outros recursos que contribuem para a melhoria de indicadores de saúde pública, como o acesso à informação.

Este estudo apresenta uma metodologia consistente para avaliar a diferença entre indicadores de desenvolvimento comparando municípios da Amazônia com outros municípios do país, contribuindo para fundamentar as considerações analíticas de especialistas para a construção de políticas públicas venham a ser implementadas, com o intuito de minimizar as desigualdades e promover o desenvolvimento da região

Amazônica no Brasil, especialmente com respeito à redução da desigualdade digital para promover o desenvolvimento da região e o acesso aos serviços básicos de saúde, educação, governo e sociedade.

REFERÊNCIAS

AGRAWAL R., IMIELŃSKI T., SWAMI A. Mining association rules between sets of items in large databases. Em: Buneman P. Jajodia S. (eds.). Sigmod '93: Proceedings of the 1993 ACM SIGMOD international conference on Management of data, pp. 207-216. New York: ACM, 1993. <http://dx.doi.org/10.1145/170035.170072>.

ATLAS: Índice de Desenvolvimento Humano do Brasil. Disponível em: <http://www.atlasbrasil.org.br/2013/pt/perfil_uf/para#caracterizacao>. Acesso em: 15 de agosto de 2020.

BRITO, S. R. Estratégia para análise da concentração de infraestrutura de acesso às tecnologias da informação e comunicação nos municípios da Amazônia legal Brasileira. 2016. 103 f. Tese (Doutorado em Engenharia Elétrica) - Universidade Federal do Pará, Belém 2016.

BRITO, S. R. et al. Concentration of access to information and communication technologies in the municipalities of the Brazilian Legal Amazon. PloS ONE, Croatia, v. 11, n. 4, p. e0152655, 2016a. BRITO, S. R. et al. Gravidez na adolescência e o acesso às Tecnologias de Informação e Comunicação na Amazônia. Mundo Amazônico, v. 6, n. 2, 2016.

CHEN, Z. Data Mining and Uncertain Reasoning: An Integrated Approach. New York: John Wiley & Sons Inc; 2001.

COOPER, G.F. E E. HERSKOVITS. 1992. A Bayesian method for the induction of probabilistic networks from data". Machine Learning 9(4): 309-347. <http://dx.doi.org/10.1007/BF00994110>

FAYYAD, U; SHAPIRO, G. P; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. AI Magazine, V. 3, 1996. p.41.

FITZ, P. R. Cartografia básica. São Paulo: Oficina de textos, 2008. 143p.

IBGE. Relatório da Divisão de Contabilidade: Relatório de Gestão. Disponível em: <<http://www.ibge.gov.br/home/presidencia/noticias/relatgestao.shtm>>. Acesso em: 25 de agosto de 2017.

ITU. International Telecommunication Union. Measuring digital development, Facts and figures, Place des Nations, Geneva Switzerland, ITU, 2019.

NERI, M. C. Mapa da Inclusão Digital. Rio de Janeiro: FGV, CPS, 2012.

NERI, M. O Mapa da Exclusão Digital. Revista Conjuntura Econômica, v. 57, n. 5, p. 70-73, 2003.

QGIS. Disponível em: <https://www.qgis.org/en/site/>.

REZENDE, S. O. Sistemas Inteligentes – Fundamentos e Aplicações. Manole (2003).

ROCHA, C. A. J. da. Redes Bayesianas para extração de conhecimento de bases de dados, considerando a incorporação de conhecimento de fundo e o tratamento de dados incompletos. Dissertação de Mestrado, ICMC-USP (1999).

SANTANA, V. V., SANTOS, P. R., LEAL, A. K. T. B., SILVA, D.B.S., PEREIRA, E.V., SILVEIRA, L.N.S., NASCIMENTO, R.A., FAGUNDES, F.E.A. (2020). A importância do uso da internet sob o viés da promoção interativa na educação em tempos de pandemia. *Brazilian Journal of Development*, v. 6, n. 8, 2020.

SANTO, R. M. DO E., & POMIN, A. V. C. O ACESSO À INTERNET COMO UM DIREITO FUNDAMENTAL. *Revista Ibero-Americana de Humanidades, Ciências e Educação*, 7(4), 399–416, 2021. <https://doi.org/10.51891/rease.v7i4.975>.

SILVA, A. S. Uma metodologia para acompanhamento e monitoramento de intervenções de formação baseada em análise e mineração de redes sociais. 2016. 105 f. Tese (Doutorado em Engenharia Elétrica) - Universidade Federal do Pará, Belém 2016.

SILVA, A. S., BRITO, S. R., BARROSO, R. F. F., CRUZ, A. G., MONTEIRO, M. A., COSTA, J. C. W. A., FRANCÊS, C. R. L. Gravidez na adolescência e associação com indicadores de renda, educação e acesso às Tecnologias de Informação e Comunicação no Brasil e na Amazônia Legal Brasileira. *Saúde e Desenvolvimento Humano. ResearchGate*, v. 4, n. 2, p. 21-33, nov. 2016.