

# Do método para estimar o tamanho médio das redes pessoais e o tamanho de populações difíceis de contar\*

Weber Soares\*\*  
Dimitri Fazito\*\*\*  
Sergio Donizete Faria\*\*\*\*

*O método conhecido como Network Scale-Up (NSU) – método de ampliação das redes sociais –, utilizado para estimar “populações difíceis de contar” (hard-to-count populations), baseia-se na ideia de que as populações humanas se organizam numa rede complexa de interações sociais, na qual todos os indivíduos, independentemente de atributos pessoais distintos, estão conectados. Conhecendo-se, então, o padrão das redes pessoais associado a determinados atributos individuais, é possível estimar “parcelas” da população que possuem esses mesmos atributos. Os emigrantes internacionais, em especial os que se encontram em situação irregular, enquadram-se nesse tipo de subpopulação, cujo tamanho é desconhecido dada a dificuldade ou até mesmo a impossibilidade de mensurá-la diretamente. A descrição do método da ampliação das redes sociais e dos procedimentos metodológicos para obtenção dos dados necessários à aplicação desse método para estimar o número de emigrantes e de retornados internacionais de uma cidade brasileira de porte médio hipotética constitui o objetivo deste texto.*

**Palavras-chave:** Método de ampliação das redes sociais. Populações difíceis de contar. Migração internacional.

## Introdução

Na escala mundial, o número de migrantes internacionais cresceu mais de 140% entre 1975 e 2005: de 82 milhões no início desse período, o total de pessoas que residiam fora de seus países de origem por mais de um ano atingiu a casa de quase 200 milhões em 2005, incluindo 9,2 milhões de refugiados (GCIM, 2005). Quanto aos fluxos migratórios irregulares, as estimativas põem à mostra que, dos 56,1 milhões de migrantes

na Europa, em 2000, pelo menos 5 milhões estavam em situação irregular (10%). Nos EUA, em 2005, cerca de 10 milhões estavam nessas mesmas condições (GCIM, 2005).

A emigração internacional de brasileiros ganha densidade em meados da década de 1980 e tem íntima conexão com a crise da economia brasileira, cujos sinais tornaram-se evidentes desde o final dos anos 1970: inflação acentuada e em ascensão; elevado endividamento externo; orçamento do governo federal onerado por subsídios

\* O presente trabalho foi realizado com o apoio de Fundação de Amparo à Pesquisa do Estado de Minas Gerais – Fapemig e do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq – Brasil.

\*\* Professor e pesquisador do Departamento de Geografia – IGC/UFMG.

\*\*\* Professor e pesquisador do Centro de Desenvolvimento e Planejamento Regional – Cedeplar/UFMG.

\*\*\*\* Professor e pesquisador do Departamento de Cartografia – IGC/UFMG.

compensatórios crescentes e pelo déficit de caixa; exacerbada concentração econômica e de renda; e acentuado desequilíbrio do balanço de pagamentos (BRUM, 1995; PEREIRA, 1985).

De acordo com as estimativas realizadas pelas embaixadas e consulados do Brasil, os emigrantes internacionais brasileiros apresentavam, no final de 2007, a seguinte distribuição territorial: na América do Norte, o estoque de brasileiros, em situação regular e irregular, seria de 1.278.650; na Europa, de 766.629; na América do Sul, de 611.708; na Ásia, de 318.285; no Oriente Médio, de 30.306; na Oceania, de 17.250; na África, de 15.323; e na América Central, de 6.611 (MRE, 2008). Os EUA constituem o país de destino preferencial desses emigrantes internacionais do Brasil: em 2007, haveria, em terras estadunidenses, cerca de 1.240.000 brasileiros residentes (MRE, 2008).

Apesar desse esforço institucional para sistematizar e organizar as informações sobre as perdas demográficas internacionais do Brasil, precisar a quantidade e o perfil dos emigrantes brasileiros no resto do mundo é algo complexo e de difícil execução: o aspecto irregular<sup>1</sup> de parcela significativa dessas perdas populacionais, a dificuldade de trabalhar dados de fontes distintas em virtude da falta de uniformidade quanto ao uso do conceito de migrante e o fato de que os quesitos sobre migração do censo brasileiro não foram, até o momento, desenhados para captar os deslocamentos internacionais de brasileiros<sup>2</sup> são elementos que desafiam a consecução de medidas mais precisas sobre tais deslocamentos. Em face desse espectro de dificuldades para quantificar, alguns pesquisadores brasileiros aperfeiçoaram métodos e técnicas destinados à estimação de saldos migratórios intercensitários, da migração de retorno e da emigração internacional (CARVALHO;

RIGOTTI, 1999; RIGOTTI, 1999; CARVALHO et al., 2001; CARVALHO; CAMPOS, 2006).

Emigrantes internacionais enquadram-se no que a literatura referente à formulação de métodos para estimar a distribuição do tamanho da rede pessoal nomeia de subpopulações difíceis de contar (*hard-to-count subpopulations*). Assim, em razão da existência de registros oficiais, pode-se obter o tamanho de certas subpopulações humanas possuidoras de algum traço comum, todavia, há outras cujo tamanho é desconhecido por causa da dificuldade ou até mesmo da impossibilidade de serem mensuradas diretamente. Entre as subpopulações difíceis de contar, estão grupos socialmente vulneráveis: os portadores de HIV; os usuários de drogas ilícitas; as vítimas de assalto; os emigrantes internacionais, em especial os que se encontram em situação irregular; etc. Não resta dúvida sobre a importância de aduzir estimativas confiáveis da quantidade de pessoas que pertencem a essas subpopulações, quando se tem em conta o campo de formulação de políticas públicas.

Grosso modo, a literatura especializada registra duas formas básicas de tratamento da informação sobre migrações internacionais que permitem auferir estimativas dos emigrantes: as técnicas diretas de mensuração dessas pessoas com base em quesitos censitários relativos à emigração internacional (VAN HOOK; BEAN, 2008; MULDER et al., 2001; COSTANZO et al., 2001; BILSBORROW et al., 1997); e as técnicas indiretas de mensuração do saldo migratório e do estoque da população emigrante internacional segundo a diferença residual entre a população observada e a esperada ao final de um período intercensitário<sup>3</sup> (CARVALHO et al., 2001; CARVALHO; CAMPOS, 2006; MULDER et al., 2001; PASSEL et al., 2004). Vale ainda registrar as recentes tentativas que, feitas nos Estados Unidos por pesquisadores do

<sup>1</sup> Conforme o Ministério das Relações Exteriores do Brasil, mais de 60% dos emigrantes brasileiros residentes nos EUA entraram, de forma irregular nesse país, por meio de redes de contatos pessoais (SOARES, 2009).

<sup>2</sup> No Brasil, o Censo 2010 deve fornecer dados mais precisos para chegar a estimativas diretas do número de emigrantes internacionais. Vale notar que o recurso às técnicas indiretas para calcular o saldo migratório internacional e o estoque de emigrantes já não pode mais ser realizado com as mesmas precisão e robustez outrora atingidas no trabalho pioneiro de Carvalho et al. (2001), pois a dinâmica demográfica brasileira mais recente não dá margem para considerar os pressupostos básicos que os procedimentos de cálculo exigem.

*Census Bureau*, conciliam diversos sistemas de dados demográficos para dar conta da população migrante, em especial dos imigrantes internacionais com *status* irregular (SCHMIDLEY; ROBINSON, 2003; PASSEL et al., 2004; COSTANZO et al., 2001).

Tendo em vista que o monitoramento das subpopulações socialmente vulneráveis e a avaliação recorrente de seu tamanho e perfil sociodemográfico são imprescindíveis para a orientação de políticas públicas (para avaliação de risco e da dinâmica das “transições”, o número de indivíduos que integram essas subpopulações é fundamental para a constituição de numeradores e denominadores de taxas diversas), vários agentes estatais têm fomentado o debate acadêmico e a pesquisa que se destina à elaboração de metodologias para estimar essas subpopulações, consoante uma relação custo-benefício relativamente equilibrada.

No âmbito do sistema das Nações Unidas, a Unaid incentivou o desenvolvimento de novas técnicas para estimar e monitorar subpopulações vulneráveis ao HIV/Aids, em especial os usuários de drogas ilícitas. Com o objetivo de estabelecer um padrão internacional para estimar as populações difíceis de contar, a Unaid realizou, em 2003, detalhado relatório sobre as metodologias mais consolidadas no campo da saúde pública, a saber: método de inquirição direta (com o uso eventual de censos ou pesquisas amostrais); método dos multiplicadores (*multiplier method*) baseado na triangulação de fontes diversas de informação; método de captura-recaptura (também com a variante de múltiplas amostras de capturas); e método de monitoramento comportamental. Nenhum desses métodos, no entanto, é universalmente aceito entre os agentes de políticas públicas e, menos ainda, entre os pesquisadores.

Diante da variabilidade das estimativas obtidas e da disputa crônica entre os proponentes desses distintos métodos, a cha-

mada metodologia de *Respondent-Driven Sampling* (RDS), baseada na ideia das redes sociais e na amostragem não probabilística do tipo “bola-de-neve” combinada à modelagem matemática de pesos e probabilidades ajustadas da amostra, tornou-se, no âmbito da saúde pública, largamente utilizada a partir de 2000 (HECKATHORN, 1997 e 2002; SALGANIK; HECKATHORN, 2004; SALGANIK, 2006). Embora tenha conquistado adeptos e se tornado predominante entre os pesquisadores e agentes de políticas em saúde pública, a metodologia de RDS está longe de ser consensual: sua maior desvantagem está no custo técnico e material elevado, o que limita, especialmente em países pobres, sua replicação.

Dificuldades dessa natureza abrem espaço para o emprego da metodologia de ampliação das redes sociais, *Network Scale-Up* (NSU), no tocante à obtenção de estimativas para subpopulações difíceis de contar. Além de permitir o emprego de pesquisa amostral probabilística, técnicas mais simples de modelagem e agentes não especializados para replicação, essa metodologia implica baixo nível de exposição dos entrevistados, ainda que o foco da pesquisa esteja voltado a subpopulações vulneráveis de grande suscetibilidade. No campo de estudos da migração, verifica-se a dificuldade técnica e metodológica de obter informações demográficas pertinentes à geração de estatísticas confiáveis sobre o tamanho e a composição das subpopulações migrantes. A aplicação do NSU para estimar os emigrantes internacionais mostra-se promissora nesse caso e toma a feição de técnica de estimativa indireta, que combina aspectos da investigação demográfica com pressupostos sociológicos sobre as interações sociais entre indivíduos e grupos possuidores de determinados atributos sociodemográficos comuns.

A descrição do método de ampliação das redes pessoais (*Network Scale-Up*

<sup>3</sup> Dos diversos vetores que compõem a dinâmica migratória, a inexistência de informações sobre o número de emigrantes internacionais impõe limites à obtenção das estimativas globais referentes aos fluxos migratórios internacionais. A aplicação das técnicas diretas e indiretas para a consecução das estimativas dos fluxos migratórios internacionais enfrenta, todavia, limitações diversas em razão da especificidade das informações censitárias que são, nesse caso, necessárias.

*Method* ou *NSU*) e das duas técnicas dele derivadas – a da estimação retroativa (*back-estimation*) e a da soma (*summation*) – constitui o eixo central da linha de exposição conduzida na primeira parte deste texto. A esse esforço descritivo serve de remate a exposição dos procedimentos da pesquisa de campo, que evidenciam os critérios de amostragem e o levantamento dos dados para estimar o número de emigrantes e de retornados internacionais de uma cidade hipotética de porte médio (66 mil domicílios): é pela aderência a tais procedimentos que se obtêm, portanto, as informações reticulares necessárias à geração de estimativas confiáveis dessas subpopulações.

### Do método e das técnicas

Segundo McCarty et al. (2001), o principal componente do método de ampliação das redes sociais, proposto e aplicado para estimar a distribuição do tamanho da rede social dos indivíduos em determinado recorte territorial e o tamanho de subpopulações difíceis de contar, é o tamanho médio das redes pessoais de uma grande amostra de indivíduos – também conhecido como parâmetro “c”.

Esse método ancora-se no pressuposto de que o número de pessoas de uma subpopulação qualquer, conhecida por um respondente, guarda relação direta com o número de pessoas que a população geral conhece dessa mesma subpopulação: o tamanho e a composição média das redes pessoais dos respondentes selecionados numa amostra são considerados representativos tanto da distribuição geral dos tamanhos médios das redes pessoais quanto da composição média (por atributos específicos) da população total. Daí, *ceteris paribus*, a probabilidade de que qualquer membro da rede de um respondente esteja em determinada subpopulação (que possui um atributo específico) deve corresponder à distribuição geral das frações de subpopulações que possuem tal atributo específico na população geral (McCARTHY et al., 2001).

Quatro conjuntos de informações são fundamentais à aplicação do método *Network Scale-Up* (NSU): tamanho médio das

redes pessoais ativas dos indivíduos da população geral (parâmetro “c”); média do total de indivíduos pertencentes às redes pessoais dos respondentes da amostra que possuam atributos específicos das subpopulações (parâmetro “m”); conjunto de subpopulações (E) com atributos específicos, cuja distribuição seja reconhecida na população total (na qual o total de cada subpopulação específica fornece o parâmetro “e”); e total observado da população geral (parâmetro “t”). À luz do pressuposto de que as associações são lineares para esse conjunto de fatores, o modelo NSU fundamenta-se na proposição de que a relação de proporcionalidade entre o total de indivíduos com um atributo específico “m” pertencentes à rede pessoal “c” do respondente amostrado deve corresponder linearmente à proporcionalidade de indivíduos “e” que possuam esse atributo da população total “t”, pois a população geral corresponde ao somatório de todos os indivíduos com seus atributos específicos. Logo, essa probabilidade ganha expressão formal na seguinte equação:

$$\frac{m}{c} = \frac{e}{t} \quad (1)$$

De acordo com McCarthy et al. (2001), os pressupostos nos quais se baseia o método são:

- todas as pessoas da população observada T possuem a mesma chance de conhecer alguém de uma subpopulação E;
- todas as pessoas da população T possuem informação perfeita sobre aqueles indivíduos que fazem parte de sua rede pessoal;
- os respondentes podem contar com precisão e em curto espaço de tempo o número de pessoas que eles conhecem (“c”) e identificar as subpopulações a que pertencem essas pessoas (“m”).

Dado que as interações sociais não são lineares, mas sim dinâmicas e pautadas em “contatos preferenciais” (BARABÁSI, 2003; SCOTT, 2000; WASSERMAN; FAUST, 1994), efeitos diversos podem causar o mascaramento ou alteração das relações esperadas

entre os parâmetros do modelo. Efeitos de barreira ocorrem quando o primeiro pressuposto é violado: existem características socio-demográficas e culturais dos respondentes e das pessoas sobre as quais eles fornecem informações que conferem ao respondente maior ou menor conhecimento a respeito de uma subpopulação específica, o que não caberia esperar no caso de uma amostra aleatória desses mesmos respondentes (McCORMICK; SALGANIK; ZHENG, 2009). Por exemplo, em região onde não exista o fenômeno da emigração internacional, não terá o respondente aí amostrado alguém que, na sua rede pessoal, pertença a essa subpopulação. Todavia, os efeitos de barreira são negligenciáveis quando a amostra de respondentes é representativa e o tamanho das subpopulações aproxima-se de 5% da população geral (McCARTHY et al., 2001). Assim, para captar uma subpopulação desconhecida e com baixa representatividade, como é o caso dos emigrantes internacionais, o modelo demanda o conhecimento do tamanho de um conjunto de subpopulações com representatividade adequada ao procedimento de ajuste da estimativa.

A violação do segundo pressuposto implica efeitos de transmissão: a informação sobre o fato de que uma pessoa pertence a certa subpopulação não é transmitida com igual probabilidade a todos que conhecem essa pessoa, em virtude do estigma associado ao pertencimento a grupos sociais específicos ou porque a informação relativa a esse mesmo pertencimento não entra no campo da conversação rotineira por ser de natureza muito pessoal (McCARTHY et al., 2001). Por exemplo, mesmo numa região de "cultura migratória" estabelecida pode ocorrer de as pessoas não terem informação completa sobre os conhecidos que emigraram, por causa do estigma da migração irregular ou simplesmente pela reduzida frequência dos contatos. Esse efeito é de difícil mensuração, pois muito pouco se sabe sobre o volume de informação que os respondentes possuem a respeito de quem eles conhecem (McCORMICK; SALGANIK; ZHENG, 2009). Alguns estudos aplicados do NSU têm buscado reduzir os erros de estimativas advindos dos efeitos de trans-

missão pelo refinamento dos estimadores, contudo sem muito sucesso até o momento (SALGANIK et al., 2011).

Os efeitos de contagem originam-se da incapacidade de os respondentes registrarem com precisão o número de pessoas que eles conhecem numa subpopulação particular. O problema, muitas vezes, encontra-se, aos olhos do respondente, na ambiguidade das fronteiras subpopulacionais, ou ainda devido aos erros de memória. Verifica-se, além disso, a tendência de o respondente subestimar o número de pessoas pertencentes a grandes subpopulações e sobre-estimar esse número no caso de subpopulações pequenas (McCORMICK; SALGANIK; ZHENG, 2009).

Vale notar que, embora o método NSU apresente as limitações expostas anteriormente, apenas os efeitos de transmissão representam de fato um problema maior, o que corresponde mais à falha de precisão do que de direção e valor pontual das estimativas. Além disso, a aplicação do método para estimar populações difíceis de contar, especialmente na área da saúde pública, tem se mostrado muito satisfatória em comparação com as estimativas oriundas de métodos tradicionais (como captura-recaptura, método multiplier, RDS, ou técnicas diretas) (SALGANIK et al., 2011).

### Técnica de estimação retroativa

Consoante a teoria de redes sociais, o maior problema na aplicação do método de ampliação das redes sociais está em encontrar uma estimativa confiável do parâmetro "c" (KILLWORTH et al., 1998b). A técnica de estimação retroativa (*back-estimation*) constitui desenvolvimento desse método de ampliação das redes e serve à consecução de estimativas confiáveis do tamanho médio das redes pessoais, isto é, do parâmetro "c".

A estimação retroativa baseia-se na reconstituição do tamanho médio das redes sociais dos indivíduos ("c"), com base no ajuste das proporções dos membros da rede pessoal do respondente que têm um conjunto de atributos ("m"). Esse conjunto guarda correspondência com o grupo observado das proporções das subpopulações

conhecidas “e” pertencentes à população geral “t”. Então, o isolamento algébrico da variável “c” na equação (1), de acordo com um modelo de máxima verossimilhança, ganha evidência na seguinte equação (KILLWORTH et al., 1998a):

$$c_i = t \cdot \frac{\sum m_{ij}}{\sum e_j} \quad (2)$$

na qual “i” representa o respondente e “j” corresponde à subpopulação conhecida.

Segundo Killworth et al. (1998a, p. 293-294), esse é um estimador não viciado porque o tamanho “j” das subpopulações “e” é relativamente pequeno se comparado à população total “t” (com uma proporção próxima ou abaixo de 5%). Como o erro padrão é dado pela raiz quadrada do produto entre “c” e “t”, divididos pelo somatório das subpopulações “e”, a acurácia do parâmetro será tanto maior quanto maior for o número de subpopulações conhecidas (E). Por isso, para assegurar a precisão do estimador “c”, recomendam Killworth et al. (1998a) que cerca de 20 subpopulações conhecidas sejam utilizadas na aplicação da técnica de estimação retroativa – com 30 subpopulações os ganhos se estabilizam (KILLWORTH et al., 1998a).

Assim, por meio dessa técnica, o tamanho da rede pessoal que maximiza a probabilidade de um tamanho médio “c” na população geral é obtido com base nos dados relativos ao número de pessoas da rede pessoal dos respondentes que integram diversas subpopulações de tamanho conhecido. De posse do tamanho médio dessas redes, uma estimativa de probabilidade máxima é computada para obter o tamanho fracionário de subpopulações desconhecidas (parâmetro “e”) (KILLWORTH et al., 1998b; McCARTHY et al., 2001).

#### *Técnica da soma*

Com o objetivo de reduzir os efeitos de transmissão e de barreira sobre as estimativas do tamanho das redes pessoais, foi elaborada a técnica da soma, também com a intenção de se chegar a uma estimativa confiável do parâmetro “c”. Uma

diferença fundamental entre a técnica de estimação retroativa e a da soma está no tipo de informação que o respondente torna disponível: em vez de fornecer a distribuição do número de pessoas integrantes da rede pessoal em várias subpopulações, o respondente é instado a contar o número de pessoas que fazem parte de diferentes categorias de relações pessoais, tais como familiares, vizinhos, colegas de trabalho, etc. (McCARTHY et al., 2001).

As vantagens potenciais do uso de diferentes categorias de relações pessoais, em vez de recorrer a subpopulações contáveis para estimar o tamanho da rede pessoal, são arroladas por McCarthy et al. (2001):

- é mais fácil para o respondente contar as pessoas segundo as categorias de relações pessoais do que pedir a ele a distribuição das pessoas da sua rede pessoal pelas subpopulações conhecidas;
- em relação à técnica de estimação retroativa, a técnica da soma fornece uma estimativa direta, sendo de mais rápida execução quando se consideram as dificuldades do respondente para situar os integrantes de sua rede pessoal em 20 ou 30 subpopulações de tamanho conhecido de acordo com o critério de pertencimento;
- a técnica de estimação retroativa depende da disponibilidade de informações precisas sobre o tamanho das subpopulações, o que muitas vezes é difícil de serem obtidas, especialmente em países com sistemas precários de informação demográfica, ou em sistemas que não possibilitam completa desagregação espacial das informações;
- o recurso à técnica da soma facilita grande redução tanto dos efeitos de transmissão – pois o respondente quase sempre sabe quem faz parte ou não do seu campo de relações particulares –, quanto dos efeitos de barreira – porque não caberia esperar impedimentos de ordem cultural e sociodemográfica para conhecer

membros de uma rede de relações particulares;

- a técnica da soma guarda independência da técnica de estimação retroativa para estimar o parâmetro “c”, o que permite utilizar as estimativas referentes às subpopulações de tamanho conhecido para verificar a precisão das estimativas sobre as subpopulações de tamanho desconhecido.

Quanto às desvantagens potenciais do uso de relações pessoais para estimar o tamanho médio das redes pessoais, McCarthy et al. (2001) assinalam:

- a inexistência de uma maneira de verificar estatisticamente a validade das estimativas referentes ao número de pessoas conhecidas pelo respondente em várias categorias relacionais, uma vez que não se sabe o tamanho real dessas categorias – além do fato de a dinâmica social impor variações complexas a cada caso;
- a frequente contagem de um mesmo membro da rede mais de uma vez, o que resulta em certa inflação do parâmetro “c” e deflação da estimativa de algumas das subpopulações desconhecidas.

Apenas o cálculo do somatório de todos os membros das redes pessoais pertencentes a todas as categorias de relações definidas em pesquisa é necessário à aplicação da técnica da soma, daí sua simplicidade. Vale mencionar que os procedimentos metodológicos adotados por Salganik et al. (2011) e por Fazito (2009), que trabalharam com 22 categorias de relações, na condução da pesquisa realizada em Curitiba, Paraná, serviram de base à replicação da técnica em outros contextos nacionais.

### Procedimentos de pesquisa

As duas técnicas apresentadas anteriormente, que permitem chegar a estimativas

confiáveis do parâmetro “c” e assim calcular o tamanho de subpopulações difíceis de contar, baseiam-se na imaginação sociológica fundamental de que as pessoas organizam suas relações pessoais segundo um padrão social recorrente, que reflete a estrutura hierárquica da realidade social empírica, isto é, as pessoas identificam e contam seus conhecidos. Além disso, elas tendem a classificar e categorizar esses conhecidos e suas interações cotidianas de acordo com uma percepção padronizada sobre a realidade imediata. Logo, é razoável admitir a existência de critérios sociais (e biológicos<sup>4</sup>) para determinar a identificação e categorização (tipificação) das pessoas e de suas relações em grupos/perfis específicos, seguindo uma hierarquia socialmente construída (FAZITO, 2009).

O padrão de identificação, contagem e classificação dos indivíduos e de suas relações numa rede de contatos demanda uma definição rigorosa das categorias de relações pessoais que, em geral, as pessoas identificam, nas quais os indivíduos incluem seus contatos pessoais cotidianos ou não e permitem aos respondentes contar e classificar efetivamente cada contato identificado. Desse modo, o processo de identificação e classificação das pessoas e de suas interações sociais pode estar fortemente associado à ideia de proximidade (por *status* social, afetividade, ou distância física), de frequência e/ou de intensidade. Assim, a forma pela qual as pessoas elaboram mentalmente as categorias sociais, identificam e contam cada indivíduo e as interações entre esses indivíduos em cada uma das categorias de relações pessoais é suscetível de reconhecimento (FAZITO, 2009).

### *Crítérios de amostragem*

O recurso a essas técnicas para estimar uma população difícil de contar – como é caso dos brasileiros oriundos de uma cidade de porte médio qualquer que residam fora do país (emigrantes internacionais) ou

<sup>4</sup> Quanto à existência de critérios biológicos para a definição de um padrão de classificação social dos indivíduos segundo a percepção de um observador (ego), ver as contribuições de Robin Dunbar (1997) que registra, como número aproximado do tamanho médio das redes sociais humanas, 150 indivíduos.

dos brasileiros que, depois de morar no estrangeiro, voltam a residir nessa cidade de origem (retornados internacionais) – implica necessariamente a realização de levantamento amostral. A amostragem deve estar claramente vinculada aos setores censitários da cidade, o que torna possível a obtenção de uma amostra probabilística dos domicílios citadinos, a qual pode ser constituída por estratificação em múltiplos estágios, desde que haja também uma amostra emparelhada para substituição (caso seja necessária durante o processo de coleta) e os domicílios substituintes pertençam ao mesmo estrato dos domicílios originais. Além disso, a amostra deve respeitar o limite mínimo de dez e máximo de 14 domicílios por setor censitário.

O conjunto de unidades domiciliares forma, então, a base da amostragem; e o conceito de domicílio, unidade de amostragem, refere-se a qualquer moradia estruturalmente independente, constituída por um ou mais cômodos, com entrada privativa, onde residem uma, duas, ou no máximo cinco famílias, mesmo que esteja localizado em estabelecimento industrial, comercial, etc. Excluem-se os domicílios coletivos, tais como hotéis, pensões, asilos, conventos, etc., pois neles a relação entre os moradores restringe-se à subordinação de ordem administrativa e ao cumprimento de normas de convivência.

A representatividade da população de domicílios é garantida, nesse caso, pela definição do tamanho mínimo da amostra de acordo com fórmula estatística para amostra aleatória simples, universo finito, nível de confiança 95% e erro amostral tolerável de 4%. A etapa inicial da rotina de cálculos consiste em obter a primeira aproximação do tamanho da amostra. Assim:

$$n_0 = \frac{1}{E_0^2} \quad (3),$$

na qual “ $n_0$ ” é a primeira aproximação do tamanho da amostra e “ $E_0$ ” corresponde ao erro amostral tolerável. De posse dessa primeira aproximação, o tamanho da amostra é obtido pela seguinte equação:

$$n = \frac{N \cdot n_0}{N + n_0} \quad (4),$$

em que “ $N$ ” é o número de elementos da população e “ $n$ ” refere-se ao tamanho da amostra.

Logo, para uma cidade de porte médio que abriga aproximadamente 66 mil domicílios, deve ser submetida à pesquisa uma amostra de 620 domicílios. Delimitadas as áreas efetivamente ocupadas em cada bairro citadino, a determinação espacial dos domicílios a serem pesquisados ancora-se em dois procedimentos: sorteio, sem reposição, de um total de ruas equivalentes ao número de domicílios amostrados para cada bairro, tendo por base o cadastro técnico municipal; e definição do lugar nas ruas selecionadas pela escolha aleatória de um percentual entre 0% e 100% (o primeiro percentual relaciona-se ao início da rua e o segundo ao seu comprimento total).

Substituições domiciliares, causadas por problemas diversos, que se tornem imperativas durante a fase de aplicação dos questionários orientam-se consoante o percentual sorteado e o lado da rua correspondente aos números pares ou ímpares, isto é, o domicílio selecionado para substituição será sempre o mais próximo à direita ou à esquerda do lugar definido em planta baixa da cidade, mediante o sorteio do sentido que deve tomar o entrevistador.

O respondente em cada domicílio, unidade de pesquisa, corresponde a qualquer residente com idade igual ou superior a 18 anos. Entre os domiciliados que pertencem a essa faixa etária, o expediente para a seleção do respondente consiste no “método do próximo aniversário”, ou seja, a pessoa que, por ocasião da pesquisa em determinado domicílio, primeiro faz aniversário entre os residentes é a que deve prestar as informações necessárias.

#### *Levantamento de dados*

A concepção de rede social que orienta a formulação do conjunto de perguntas (Quadros 1 e 2) destinado à obtenção dos parâmetros “ $m$ ” e “ $c$ ” é o de rede ativa, ou seja, essa rede se constitui de pessoas que o respondente conhece de vista ou de nome, pessoas essas que conhecem o respondente de vista ou de nome e com



as quais ele pode entrar em contato caso queira.

Quanto às questões referentes ao identificador de subpopulações (“m”), os respondentes são inqueridos a fornecer o número de pessoas que eles conhecem numa subpopulação específica “k”. Essas questões (Quadro 1), empregadas para estimar o tamanho da subpopulação de emigrantes internacionais de uma cidade de porte médio qualquer e de retornados internacionais a essa mesma cidade, classificam-se em dois subconjuntos. O primeiro, com 23 questões, corresponde a subpopulações de tamanho conhecido. Vale notar que as questões referentes às subpopulações de tamanho conhecido discriminam-se, ainda, de acordo com a natureza do impacto, em questões de baixa sensibilidade e de alta sensibilidade. Essa distinção apoia-se em procedimento adotado por Snidero et al. (2010) que considera estas últimas como questões que mais diretamente sensibilizam o respondente, tal como a morte ou a doença de um parente ou de um amigo.

O segundo subconjunto guarda pertinência com as subpopulações de tamanho desconhecido e abarca sete questões-alvo (Quadro 1). Trata-se de informações sobre o número de emigrantes e retornados internacionais integrantes da rede pessoal do respondente, que permitem estimar o tamanho de cada uma dessas subpopulações.

Constam, no Quadro 2, as questões que servem à obtenção do tamanho médio das redes pessoais, parâmetro “c”. De acordo com McCarthy (2001), a definição das categorias de relações pessoais deve ser adaptada ao contexto sociocultural dos respondentes. No caso do Brasil, em pesquisas anteriores, foi definido o conjunto de 22 categorias de relações apresentadas no Quadro 2 (FAZITO, 2009; SALGANIK et al., 2011). Elas são elaboradas com base em três categorias gerais (familiares, amigos e conhecidos casuais) clivadas e subdivididas em outras subcategorias definidas por eixos transversais específicos (proximidade física, graus de intimidade e tipos de atividade). Ademais, para que o respondente possa fornecer/contar o número exato de pessoas que fazem parte de sua rede social ativa, e

não apenas indicar um número aproximado, as categorias de relações pessoais devem ser exaustivas, mutuamente exclusivas e suficientemente pequenas, devendo seguir uma hierarquia de inclusão, do ponto mais estrito e íntimo no círculo social até o mais abrangente e casual dos contatos pessoais (pode-se imaginar algo como um círculo se ampliando e se tornando mais inclusivo a partir do indivíduo – ego), além de que cada categoria deve ser pequena o suficiente para o respondente poder contar efetivamente cada pessoa constante no seu mapa mental de contatos pessoais (FAZITO, 2009).

## Discussão

O modelo de ampliação das redes pessoais e a técnica de estimação retroativa, que daquele constitui extensão, foram empregados em estudo-piloto em Governador Valadares, cidade reconhecida por sua “cultura migratória” internacional, para obter as estimativas referentes ao tamanho médio das redes pessoais e ao total de emigrantes internacionais. Assim, para essa cidade, composta de aproximadamente 66 mil domicílios, as informações relativas a uma amostra de 646 domicílios – que em correspondência com o nível de confiança de 95% e erro amostral tolerável de 4% garante a representatividade das informações relativas ao identificador de subpopulação (número de emigrantes internacionais dessa cidade) e aos geradores de posição (tamanho da rede pessoal) – possibilitaram as estimativas que se seguem: 194 pessoas expressaram o tamanho médio da rede pessoal valadarense; o total de emigrantes internacionais correspondeu a 6.642 pessoas. Esta última estimativa refere-se ao estoque de emigrantes internacionais num período específico, apenas um ano, e leva necessariamente em conta indivíduos que mantêm interações sociais ativas entre origem e destino.

Pelos cálculos, Valadares apresenta proporção relativamente robusta de emigrantes internacionais: 2,6% de sua população possui um atributo raro: ser emigrante internacional, o que mostra grande conformidade com os diversos estudos sobre

**QUADRO 1**  
**Questões relativas a subpopulações de tamanho conhecido e de tamanho desconhecido**

Identificador de subpopulações – parâmetro “m”		
Subpopulações de tamanho conhecido		Tipo de questão
1.	Das mulheres que você conhece e vivem nesta cidade, quantas tiveram filhos nos últimos 12 meses?	Baixa sensibilidade
2.	Das pessoas que você conhece e vivem nesta cidade, quantas se casaram nos últimos 12 meses?	
3.	Das pessoas que você conhece e vivem nesta cidade, quantas são pais ou mães de três ou mais de três filhos?	
4.	Das pessoas que você conhece e vivem nesta cidade, quantas trabalham atualmente nos Correios (carteiro)?	
5.	Das pessoas que você conhece e vivem nesta cidade, quantas trabalham atualmente na prefeitura municipal?	
6.	Das pessoas que você conhece e vivem nesta cidade, quantas são motoristas de táxi?	
7.	Das pessoas que você conhece e vivem nesta cidade, quantas são motoristas de ônibus municipal?	
8.	Das pessoas que você conhece e vivem nesta cidade, quantas são cobradores de passagem de ônibus municipal (trocadores)?	
9.	Das pessoas que você conhece e vivem nesta cidade, quantas são professoras(es) do ensino fundamental?	
10.	Das pessoas que você conhece e vivem nesta cidade, quantas são professoras(es) do ensino superior?	
11.	Das pessoas que você conhece e vivem nesta cidade, quantas são mulheres aposentadas?	
12.	Das pessoas que você conhece e vivem nesta cidade, quantas são homens aposentados?	
13.	Das pessoas que você conhece e vivem nesta cidade, quantas são aposentadas por invalidez?	
14.	Das pessoas que você conhece e vivem nesta cidade, quantas são evangélicas?	
15.	Das pessoas que você conhece e vivem nesta cidade, quantas são católicas?	
16.	Das pessoas que você conhecia e viviam nesta cidade, quantas morreram assassinadas nos últimos 12 meses?	
17.	Das pessoas que você conhecia e viviam nesta cidade, quantas cometeram suicídio nos últimos 12 meses?	
18.	Das pessoas que você conhecia e viviam nesta cidade, quantas morreram de acidente de trânsito nos últimos 12 meses?	
19.	Das pessoas que você conhece e vivem nesta cidade, quantas foram vítimas de sequestro nos últimos 12 meses?	
20.	Das pessoas que você conhece e vivem nesta cidade, quantas são viúvas?	
21.	Das pessoas que você conhece e vivem nesta cidade, quantas são viúvos?	
22.	Das pessoas que você conhece e vivem nesta cidade, quantas contraíram HIV/Aids?	
23.	Das pessoas que você conhece e vivem nesta cidade, quantas são diabéticas?	
24.	Das pessoas que você conhece e viviam nesta cidade, quantas moram/trabalham atualmente em outro país?	Alvo
25.	Entre essas pessoas que moram/trabalham atualmente em outro país, quantas são mulheres e quantas são homens?	
26.	Entre essas pessoas que moram/trabalham atualmente em outro país, quantas abriram algum tipo de negócio nesta cidade?	
27.	Das pessoas que você conhece, que viviam nesta cidade e foram morar/trabalhar na Itália, quantas retornaram e atualmente residem nesta cidade?	
28.	Entre essas pessoas que retornaram da Itália e vivem atualmente nesta cidade, quantas são mulheres e quantas são homens?	
29.	Das pessoas que você conhece, que viviam nesta cidade e foram morar/trabalhar nos Estados Unidos, quantas retornaram e atualmente residem nesta cidade?	
30.	Entre essas pessoas que retornaram dos Estados Unidos e vivem atualmente nesta cidade, quantas são mulheres e quantas são homens?	

Fonte: Killworth et al. (1998); McCarty et al. (2001); Zheng, Salganik e Gelman (2006); Snidero et al. (2009); McCormick, Salganik e Zheng (2009); e Fazito (2009).

**QUADRO 2**  
**Questões relativas às redes pessoais**

<b>Gerador de posições – parâmetro “c”</b>	
1.	Sem contar com você, quantas pessoas fazem parte da sua família mais íntima e vivem no seu domicílio? Conte todos os seus familiares consanguíneos e não consanguíneos (por exemplo, mãe, pai, irmão, incluindo madrasta, padrasto, enteada, etc.).
2.	Quantas pessoas fazem parte da sua família mais íntima e vivem em domicílios diferentes do seu, mas apenas nesta cidade? Conte todos os familiares consanguíneos e não consanguíneos que você considera parte da sua família próxima, mas não conte ninguém que já tenha sido contado anteriormente.
3.	Quantas pessoas fazem parte da sua família no lado paterno que são menos íntimas e moram nesta cidade? Não conte quem já foi incluído nas categorias anteriores.
4.	Quantas pessoas fazem parte da sua família no lado materno que são menos íntimas e moram nesta cidade? Não conte quem já foi incluído nas categorias anteriores.
5.	Quantas pessoas fazem parte da sua família agregada do lado do(a) seu (sua) esposo(a) ou namorado(a) e moram nesta cidade? Não conte quem já foi incluído nas categorias anteriores.
6.	Quantas pessoas você considera amigas/amigos íntimas? Não conte nenhum membro de sua família; conte todos os amigos íntimos que residem, atualmente, nesta cidade.
7.	Quantas pessoas você considera amigas do dia a dia com quem você se encontra regularmente na vizinhança da sua residência?
8.	Quantas pessoas você considera amigas do dia a dia com quem você se encontra normalmente no seu ambiente de trabalho?
9.	Quantas pessoas você considera amigas do dia a dia com quem você se encontra normalmente no seu ambiente de estudo?
10.	Quantas pessoas você considera amigas do dia a dia com quem você se encontra normalmente durante suas atividades sociais de lazer (recreação, festas, atividades físicas, etc.)?
11.	Quantas pessoas você considera amigas do dia a dia com quem você se encontra normalmente durante suas atividades religiosas?
12.	Quantas pessoas você considera amigas do dia a dia com quem você se encontra normalmente durante suas atividades cívicas (clubes, associações partidárias, voluntariado, etc.)?
13.	Quantas pessoas você considera amigas do dia a dia com quem você se relaciona através de seus familiares e encontra regularmente?
14.	Quantas pessoas você considera amigas do dia a dia com quem você se relaciona através de outros amigos e encontra regularmente?
15.	Quantas pessoas você considera conhecidos casuais com quem você se encontra geralmente na vizinhança de sua residência?
16.	Quantas pessoas você considera conhecidos casuais com quem você se encontra geralmente no seu ambiente de trabalho?
17.	Quantas pessoas você considera conhecidos casuais com quem você se encontra geralmente no seu ambiente de estudo?
18.	Quantas pessoas você considera conhecidos casuais com quem você se encontra geralmente durante suas atividades sociais de lazer (recreação, festas, atividades físicas, etc.)?
19.	Quantas pessoas você considera conhecidos casuais com quem você se encontra geralmente durante suas atividades religiosas?
20.	Quantas pessoas você considera conhecidos casuais com quem você se encontra geralmente durante suas atividades cívicas (clubes, associações partidárias, voluntariado, etc.)?
21.	Quantas pessoas você considera conhecidos casuais que você conheceu através dos seus familiares e encontra geralmente?
22.	Quantas pessoas você considera conhecidos casuais que você conheceu através de outros amigos e encontra geralmente?

Fonte: Fazito (2009) e Salganik et al. (2011).

populações difíceis de contar (KILLWORTH et al., 1998a e b; SALGANIK et al., 2011). Vale notar que, provavelmente, essa subpopulação migrante deve estar subestimada em alguma ordem por motivos diversos, tais como erros de memória, negação da informação referente ao emigrante e “relativização” do *status* do emigrante.

## Conclusão

Submetidos aqui à descrição, o método de ampliação das redes pessoais e as duas técnicas que dele constituem extensão, a da estimação retroativa e a da soma, consagram-se ao principal objetivo de oferecer estimativas confiáveis do tamanho médio das redes pessoais. É com base nessa informação de natureza reticular que o tamanho de determinada subpopulação difícil de contar pode ser obtido.

Lastreado no princípio básico de que as redes sociais são, em média, representativas da população, o método da ampliação das redes sociais demanda, na sua consecução, recurso a uma amostra aleatória da população geral. Assim, para uma cidade hipotética de 66 mil domicílios,

a amostra de 620 domicílios, com nível de confiança de 95% e erro amostral tolerável de 4%, garante a representatividade das informações relativas aos identificadores de subpopulações (número de retornados e de emigrantes internacionais dessa cidade) e aos geradores de posição (tamanho da rede pessoal).

Ainda que o modelo esteja sujeito a certos problemas (os efeitos de barreira podem acentuar a variância e conduzir a certos vieses se o quadro amostral for incompleto; os respondentes podem ser incapazes ou se recusarem a responder com precisão a certos quesitos relacionais; e falta de procedimentos seguros para aplicar intervalos de confiança às estimativas), muito esforço no âmbito da pesquisa teórica tem sido feito para reduzir a variância e magnitude desses vieses e garantir maior aderência entre o modelo e as instâncias do real que ele busca apreender.

Enfim, a pesquisa no campo das redes sociais está, de fato, apenas começando e muito resta por investigar sobre a estrutura das redes pessoais e a variação nas propensões individuais para formar laços com as pessoas de certos grupos sociais.

## Referências

BARABÁSI, A-L. **Linked – how everything is connected to everything else and what it means for business, science and everyday life**. New York: Plume Penguin Books, 2003.

BILSBORROW, R. E.; HUGO, G.; OBERAI, A. S.; ZLOTNIK, H. **International migration statistics: guidelines for improving data collection systems**. Genebra: UNFPA, International Labour Office, 1997.

BRUM, A. F. **Desenvolvimento econômico brasileiro**. 15. ed. Petrópolis: Vozes, 1995.

CARVALHO, J. A. M.; MAGALHÃES, M. V.; GARCIA, R. A.; SOARES, W. Sinuosos caminhos para estimação dos emigrantes internacionais de 1986/1991 e de 1991/1996 e saldos migratórios dos quinquênios entre 1981 e 1996 das Unidades da Federação Brasileira. In: ENCONTRO NACIONAL DE

ESTUDOS POPULACIONAIS, 12., 2000, Caxambu. **Brasil 500 anos: mudanças e continuidades**. Campinas: Abep, 2000. (Disponível em CD-ROM).

CARVALHO, J. A. M. de; CAMPOS, M. B. A variação do saldo migratório internacional do Brasil. **Estudos Avançados**, São Paulo, v. 20, n. 57, p. 55-58, maio/agosto 2006.

CARVALHO, J. A. M.; MAGALHÃES, M. V.; GARCIA, R. A.; SOARES, W. Estimativas dos saldos migratórios internacionais e do número de emigrantes internacionais das grandes regiões do Brasil – 1986/1991 e 1991/1996. In: CASTRO, M, G. **Migrações internacionais: contribuições para políticas**. Brasília: CNPD, 2001, p. 243-252.

CARVALHO, J. A. M.; RIGOTTI, J. I. R. Os dados censitários sobre migrações internas: algumas sugestões para a análise. **Revista**

**Brasileira de Estudos de População**, Campinas, v. 15, n. 2, p. 7-17, jul./dez. 1999.

COSTANZO, J. et al. **Evaluating components of international migration: the residual foreign-born**. Washington D.C.: Population Division, US Census Bureau, p. 1-24, 2001 (Working paper series, 61).

FAZITO, D.; RIOS-NETO, E. L. G. Emigração internacional de brasileiros para os Estados Unidos: as redes sociais e o papel de intermediação nos deslocamentos exercido pelas agências de turismo. **Revista Brasileira de Estudos de População**, Campinas, v. 25, n. 2, p. 305-323, jul./dez. 2008.

FAZITO, D. **Aplicação da técnica de ampliação da rede social para estimar o tamanho da população de usuários de drogas ilícitas em Curitiba, Brasil**. Belo Horizonte. 2009. Protocolo de pesquisa. Mimeografado.

GLOBAL COMMISSION ON INTERNATIONAL MIGRATION (GCIM). **Migration in an interconnected world: new directions for action**. Report of the Global Commission on International Migration, 2005. Disponível em: <<http://www.gcim.org/en/>>. Acesso em: 20 jun. 2007.

HECKATHORN, D. D. Respondent-driven sampling: A new approach to the study of hidden populations. **Social Problems**, v. 44, n. 2, p. 174-199, 1997.

\_\_\_\_\_. Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. **Social Problems**, v. 49, n. 1, p. 11-34, 2002.

HILL, R. A.; DUNBAR, R. I. M. Social network size in humans. **Human Nature**, New York, v. 14, n. 1, p. 53-72, 2003.

KILLWORTH, P. D.; McCARTY, C.; BERNARD, H. R.; SHELLY, G. A.; JOHNSEN, E. C. Estimation of seroprevalence, rape, and homelessness in the U.S. using a social network approach. **Evaluation Review**, v. 22, p. 289-308, 1998a.

KILLWORTH, P. D.; JOHNSEN, E. C.; McCARTY, C.; SHELLY, G. A.; BERNARD, H.

R. A social network approach to estimating seroprevalence in the United States. **Social Networks**, Elsevier, v. 20, p. 23 -50, 1998b.

McCARTY, C.; KILLWORTH, P. D.; BERNARD, H. R.; JOHNSEN, E. C.; SHELLY, G. A. Comparing two methods for estimating network size. **Human Organization**, Oklahoma, v. 60, n. 1, p. 28-39, 2001.

McCORMICK, T. H.; SALGANIK, M. J.; ZHENG, T. **How many people do you know?** Efficiently estimating personal network size. New York: Columbia University, 2009. Disponível em: <[http://www.stat.columbia.edu/~cook/movabletype/archives/2008/10/how\\_many\\_people\\_1.html](http://www.stat.columbia.edu/~cook/movabletype/archives/2008/10/how_many_people_1.html)>. Acesso em: 3 dez. 2009.

MINISTÉRIO DAS RELAÇÕES EXTERIORES (MRE). **Brasileiros no mundo**. Brasília: Departamento Consular e de Brasileiros no Exterior, 2008. Disponível em: <[http://www.abe.mre.gov.br/mundo/america-do-sul/republica-federativa-do-brasil/subsecretaria-geral-das-comunidades-brasileiras-no-exterior/avisos/newsitem\\_view2?id=brasileiros-no-mundo](http://www.abe.mre.gov.br/mundo/america-do-sul/republica-federativa-do-brasil/subsecretaria-geral-das-comunidades-brasileiras-no-exterior/avisos/newsitem_view2?id=brasileiros-no-mundo)>. Acesso em: 21 ago. 2009.

MULDER, T. J. et al. U.S. **Census Bureau measurement of net international migration to the United States: 1999-2000**. Washington D.C.: Population Division, Census Bureau, p. 1-24, 2001 (Working paper series, 51).

PASSEL, J. S.; VAN HOOK, J.; BEAN, F. D. **Estimates of the legal and unauthorized foreign-born population for the United States and selected states, based on Census 2000**. Final Draft. California: SABRE Systems Inc., 2004, p. 1-46.

PEREIRA, L. C. B. **Desenvolvimento e crise no Brasil: 1930-1983**. São Paulo: Brasiliense, 1985.

PORTES, A. Economic sociology and the sociology of immigration: a conceptual overview. In: PORTES, A. (Org.). **The economic sociology of immigration: essays on networks, ethnicity, and entrepreneurship**. New York: Russel Sage Foundation, 1995, p. 1-41.

RIGOTTI, J. I. R. **Técnicas de mensuração das migrações, a partir de dados censitários**: aplicação aos casos de Minas Gerais e São Paulo. 1999. Tese (Doutorado em Demografia) – Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, 1999.

SALGANIK, M. J. Variance estimation, design effects, and sample size calculations for respondent-driven sampling. **Journal of Urban Health**, v. 83, n. 7, p. 98-112, 2006.

SALGANIK, M. J.; HECKATHORN, D. D. Sampling and estimation in hidden populations using respondent-driven sampling. **Sociological Methodology**, v. 34, p. 193-239, 2004.

SALGANIK, M. J.; MELLO, M. B.; ABDO, A. H.; BERTONI, N.; FAZITO, D.; BASTOS, F. I. The game of contacts: estimating the social visibility of groups. **Social Networks**, v. 33, p. 70-78, 2011.

SCHMIDLEY, A. D.; ROBINSON, J. G. **Measuring the foreign-born population in the United States with the Current Population Survey: 1994-2002**. Washington: D.C.: Population Division, US Census Bureau, 2003.

SCOTT, J. **Social network analysis** – a handbook. London, UK: Sage Publications, 2000.

SNIDERO, S.; ZOBEC, F.; BERCHIALLA, P.; CORRADETTI, R.; GREGORI, D. Question order and interviewer effects in CATI Scale-up Survey. **Surveys Sociological Methods & Research**, v. 38, p. 287-305, 2009.

SOARES, W. Retornados internacionais e os intermediários da rede migratória valadarensense. **Revista Interdisciplinar da Mobilidade Humana – REMHU**, Brasília, Ano XVII, n. 32, p. 47-58, 2009.

UNAIDS. **Estimating the size of populations at risk for HIV**. Geneva, n. 03.36E, 2003.

VAN HOOK, J.; BEAN, F. D. **Methodology and detailed estimates of foreign-born emigration, return immigration, and net emigration, 1996-2005**. California: Sabre Systems Inc., 2008, p. 1-68.

WASSERMAN, S.; FAUST, K. **Social network analysis** – methods and applications. Cambridge, UK: Cambridge University Press, 1994.

ZHENG, T.; SALGANIK, M. J.; GELMAN, A. How many people do you know in prison?: using overdispersion in count data to estimate social structure in networks. **Journal of the American Statistical Association**, Alexandria, v. 101, n. 474, p. 409-423, Jun. 2006.

## Resumen

*Sobre el método para estimar el tamaño medio de las redes personales y el tamaño de poblaciones difíciles de calcular*

El método conocido como Network Scale-Up (NSU) –método de ampliación de las redes sociales-, utilizado para estimar “poblaciones difíciles de calcular” (hard-to-count populations), se basa en la idea de que las poblaciones humanas se organizan en una red compleja de interacciones sociales, en la que todos los individuos, independientemente de sus atributos personales diferenciados, están conectados. Conociéndose entonces el patrón de las redes personales, asociado a determinados atributos individuales, es posible estimar “porciones” de la población que posean estos mismos atributos. Los emigrantes internacionales, en especial los que se encuentran en situación irregular, se encuadran en este tipo de subpoblación, cuyo tamaño es desconocido, dada la dificultad o hasta incluso la imposibilidad de calcularlo directamente. La descripción del método de la ampliación de las redes sociales, y de los procedimientos metodológicos con el objeto de obtener los datos necesarios para la aplicación

de este método en la estimación del número de emigrantes y de retornados internacionales de una hipotética ciudad brasileña de porte medio, constituye el objetivo de este trabajo.

**Palabras-clave:** Método de ampliación de las redes sociales. Poblaciones difíciles de calcular. Migración internacional.

### **Abstract**

*Of a method for estimating the average population size of personal networks and the size of hard-to-count populations*

The method known as Network Scale-Up (NSU) – used for expanding social networks – is used to estimate hard-to-count populations, based on the idea that human populations are organized in a complex web of social interactions, to which all individuals, regardless of specific personal attributes, are connected. If we know the pattern of personal networks associated with certain individual attributes, we can estimate “parcels” of the population that have these same attributes. International migrants, especially those who are undocumented, fit into this type of subpopulation whose size is unknown because of the difficulty or even impossibility of measuring it directly. This article aims at describing the Network Scale-Up method and the methodological procedures necessary to estimate the number of international and returned migrants in a hypothetical, medium-sized, Brazilian city.

**Keywords:** Network scale-up method (NSU). Hard-to-count populations. International migration.

Recebido para publicação em 04/05/2011

Aceito para publicação em 10/12/2011