



HAL
open science

De la théorie à la pratique : quelques enseignements de la mise en oeuvre de la méthode “ Network Sampling with Memory ” pour enquêter auprès des immigrants chinois en Ile-de-France

Marine Bondon, Géraldine Charrance, Paul Cochet, Thomas Merly-Alpa, Aurélie Santos

► To cite this version:

Marine Bondon, Géraldine Charrance, Paul Cochet, Thomas Merly-Alpa, Aurélie Santos. De la théorie à la pratique : quelques enseignements de la mise en oeuvre de la méthode “ Network Sampling with Memory ” pour enquêter auprès des immigrants chinois en Ile-de-France. 14e édition des Journées de méthodologie statistique de l’Insee (JMS 2022), Mar 2022, Paris, France. hal-03922473

HAL Id: hal-03922473

<https://hal.science/hal-03922473>

Submitted on 4 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**DE LA THÉORIE A LA PRATIQUE : QUELQUES ENSEIGNEMENTS DE LA MISE EN
ŒUVRE DE LA MÉTHODE « NETWORK SAMPLING WITH MEMORY » POUR
ENQUÊTER AUPRÈS DES IMMIGRÉS CHINOIS EN ÎLE-DE-FRANCE**

*Marine BONDON, Géraldine CHARRANCE, Paul COCHET, Thomas MERLY-ALPA, Aurélie
SANTOS*

Ined, Service des enquêtes et Sondages

Mots-clés : Échantillonnage, collecte, population difficile à joindre

Domaine concerné : « Collecte de données d'enquêtes », « Protocole, conception des enquêtes, couverture de populations particulières »

Résumé

La méthode d'échantillonnage « Network sampling with memory » (NSM) constitue une nouvelle variante de sondage par chaînage (boule de neige, Respondant Driven Sampling (RDS)...). Ces méthodes consistent à enquêter au sein de réseaux en sélectionnant au départ quelques individus appelés « graines ». Seuls ces individus sont désignés par le sondeur. Par la suite, ce sont les enquêtés eux-mêmes qui recrutent/désignent leurs pairs qui seront sollicités à leur tour pour participer à l'enquête. Les méthodes par chaînage permettent en théorie d'atteindre des pans de la population non directement accessibles à des enquêteurs, mais souffrent d'un biais de sélection très fort. Dans le but de pallier ce défaut, une équipe de l'université de Caroline du Nord a développé la méthode NSM qui présente des avantages théoriques notamment sur la précision des estimations, au prix d'une complexité plus importante. La particularité de NSM est de recréer, au fur et à mesure du terrain, une base de sondage de la population cible composée des personnes citées par les répondants et de tirer aléatoirement les futurs enquêtés dans cette base. Contrairement à la méthode RDS, elle ne cherche pas à enquêter tous les contacts cités mais vise à intégrer une dimension aléatoire dans une méthode de sondage empirique. L'algorithme comprend une première phase exploratoire dite « Search », à la recherche des divers pans du réseau, puis une seconde phase de tirages aléatoires dans le réseau en prenant en compte la structure découverte en première étape. Selon ses concepteurs, la méthode NSM devrait permettre d'obtenir des estimations d'une précision équivalente à celle d'un sondage aléatoire simple.

Après deux expériences à l'étranger (en Tanzanie et en Caroline du Nord), la méthode a été utilisée pour la première fois en France entre septembre 2020 et juin 2021 dans le cadre d'une enquête menée auprès des immigrants chinois en Ile-de-France (ChIPRe). A l'issue du terrain, 500 questionnaires et quelque 1700 citations ont été collectés. Cette enquête s'est avérée extrêmement délicate à mettre en œuvre. Sa dimension aléatoire, et notamment la

nécessité de réaliser des tirages hebdomadaires a engendré de nombreuses difficultés sur le terrain. La charge de travail individuelle des enquêteurs s'est révélée difficile à anticiper, puisqu'une forme de filiation citant/cités (conserver le même enquêteur pour un enquêté et ses amis échantillonnés) était privilégiée afin de faciliter les négociations des enquêteurs sur le terrain. De plus, un paradoxe inhérent à la logique de l'algorithme de tirage dans sa phase « exploratoire » conduisait les enquêteurs qui collectaient le plus de citations (les plus grands « rosters ») à se voir attribuer parfois moins d'échantillonnés que d'autres enquêteurs moins performants. En effet, les petits rosters étaient plus volontiers candidats à l'échantillonnage car interprétés par l'algorithme comme une opportunité d'aller vers des pans inexplorés du réseau, à l'inverse des plus grands rosters, plus souvent délaissés par l'algorithme car associés à la découverte de personnes citées plusieurs fois (doublons), synonymes d'un niveau d'exploration avancé du réseau. Ainsi les « bons » enquêteurs ne voyaient pas toujours leurs efforts récompensés par le tirage des rosters qu'ils avaient collectés. Cette impossibilité d'anticiper les tirages s'est aussi concrétisée dans le travail de communication auprès des enquêtés, car les enquêteurs étaient incapables de prédire dans leurs argumentaires si les amis cités par leurs enquêtés seraient finalement tirés au sort ou non, ni à quelle échéance.

À ces difficultés se sont ajoutées celles de la pandémie. La crise sanitaire et la suspension de la collecte en face-à-face induite par le second confinement ont eu des impacts conséquents sur la collecte. Initialement pensée pour être conduite en face-à-face, nous avons été contraints de mener au moins une partie de l'enquête par téléphone. Les méthodes d'échantillonnage par chaînage reposant sur un lien de confiance entre enquêteur et enquêté, contribuant à une réelle implication des enquêtés dans le processus de recrutement, la prise de contact et l'interrogation à distance (par téléphone) a rendu la tâche plus ardue. Le distanciel a également rendu plus complexe et moins efficace la gestion des incitations financières (remise de chèques cadeaux).

Bibliographie

- [1] Mouw T, Verdery AM. « Network Sampling with Memory: A proposal for more efficient sampling from social networks». *Sociol Methodol.* 2012 Aug;42(1):206-256.
- [2] Goel S, Salganik M.J, «Assessing respondent-driven sampling», *PNAS* April 13, 2010 107 (15) 6743-6747
- [3] Merli M G, Verdery A., Mouw T, Li J., «Sampling Migrants from their Social Networks: The Demography and Social Organization of Chinese Migrants in Dar es Salaam, Tanzania», *Migr Stud.* 2016 Jul;4(2):182-214.
- [4] Mouw T, Chavez S, Edelblute H, Verdery A., «Binational Social Networks and Assimilation: A Test of the Importance of Transnationalism», *Social Problems* Vol. 61, No. 3 (August 2014), pp. 329-359
- [5] Merli G, Mouw T, Stolte A, Le Barbenchon C, Florey-Eischen F, Using Multiple Modes of Data Collection to Recruit Migrant Samples With Network Sampling With Memory: The Chinese Immigrants in Raleigh-Durham (ChIRDU) Study, 2019 Annual Meeting, Austin, TX April 10-13