CGAP

BRIEF

Can Digital Footprints Lead to Greater Financial Inclusion?

Whether in Mexico, Ghana, or Pakistan, millions of poor people wake up to the glowing light of their mobile phones. Poor people have mobile phones, but not formal financial services. CGAP and GSMA estimated that close to 2 billion people will have a mobile phone but not a bank account in 2012.

Poor people's use of their mobile phones generates data that leave what can be called a digital footprint. These data are among a handful of sources of electronic information that exist on poor people. This information is potentially powerful but has not yet been used in ways to radically impact financial access for poor people.

This Brief highlights some early experience on the potential of digital footprints from mobile phone use. Most of this initial work is experimental. As long as consumer interests are protected and privacy, security, and ethical use concerns are addressed, these data may become a useful way to reach unbanked poor people with a range of financial products.

Data behind digital footprints

There are four basic kinds of data generated by mobile phone use:

1. Timing, location, and duration of voice and text-message and airtime purchase: Mobile phone use for voice calls and text messages generates a call detail record (CDR) that is recorded by mobile network operators (MNOs) to accurately bill customers.
2. Use of value-added services, such as ringtones, text-messaging-based services: People also download ring tones, play games, subscribe to text-messaging-based information services (e.g., sport scores, agriculture pricing, health alerts, etc.) and respond to text-messaging-based surveys.
3. Internet use: While the poor have traditionally not used the Internet much, the rate of usage among those living at or near the poverty line is expected to rise as it becomes cheaper and easier to access Internet services.[1]
4. Financial transactional data: As the volume of mobile money transactions continues to grow, providers will have access to a deeper well of data.

CDR and other basic data are passive data, whereas data provided by people in response to surveys, on social media sites, and so on, are considered active data. While the term active data more accurately describes the person who generated the data, both passive and active data have potential value. For instance, Massachusetts Institute of Technology researchers were able to predict with 74 percent accuracy the sex of mobile phone users based primarily on an analysis of CDR data (Eagle et al. 2009).

The duration for which these data are available impacts the quality of the analysis. For instance, according to Wired (2001), the four major U.S. MNOs keep basic CDR data anywhere from four months to three years or longer. Regulators may or may not specify "data retention" rules. In Pakistan, carriers are required by the telecommunications regulator to keep data for at least three years.[2] In addition, some operators do not attach a new subscriber identity module (SIM) card to an actual subscriber name, making it difficult to track account-level relationships with a given customer, especially someone with two SIM cards.

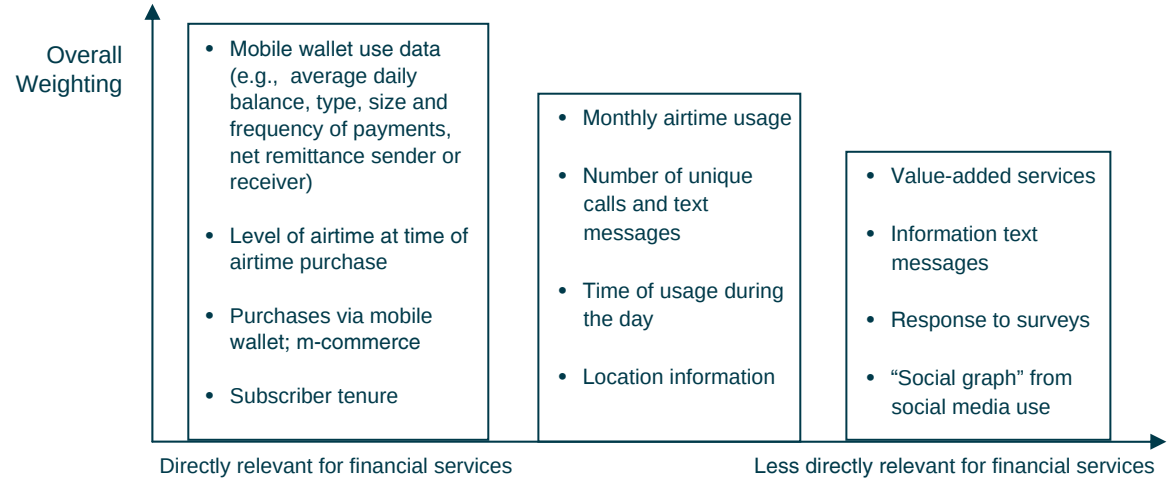
As shown in Figure 1, the variables from these data can be placed on a continuum from direct financial transaction data and other data that map a user's use of money, to data that is nonfinancial and less obviously related to financial services, but could

1 In Kenya, 90 percent of Internet usage occurs via cell phones, and 31 percent of Internet users are on Facebook. In June 2012, Airtel Ghana and Facebook launched Facebook's text-only mobile site called "Facebook Zero," which allows free access to Facebook.

2 See http://www.pta.gov.pk/images/stories/kashif/apc_rules.pdf. The data include "all books and accounts pertaining to payments made or received…and the telecommunication services to which such payments relate, including call detail records and itemized billing data."

July 2012

Figure 1: Value of variables for financial services



be insightful. The challenge is determining how to weight the variables to build an accurate profile of the customer relevant for financial services.

Use of digital footprints for financial services

In partnership with an MNO, CGAP set out to test some hypotheses on how digital footprints could be used to deliver credit to the unbanked. For example, we hypothesized that people who purchased airtime frequently and in a consistent pattern (similar amount or at similar times) demonstrated predictability in income and better planning, which might impact their ability to repay a loan. We also hypothesized that people who have an inactive prepaid account or one that consistently runs to zero airtime balance before their next airtime purchase may not be strong planners.

While no one variable in isolation is likely to be adequate to profile a customer's credit risk, our premise was that certain variables in combination might be able to do so. For example, a customer who has been active for three years, reloads the same amount of airtime minutes every Friday, and rarely lets his prepaid balance run down to zero could be rated "low risk," while a customer who had just activated his SIM and reloaded a low number of minutes now followed by periods of inactivity might be rated "high risk."

In the end, CGAP was unable to test the model, but companies such as Cignifi and Experian have gone further. Cignifi built a credit scoring model using CDR data and tested it in Tanzania and Brazil. It built a model in Brazil using 50 variables from 2.3 million prepaid customers of MNO Oi's mobile business and back-tested the model against historical lending data from approximately 40,000 borrowers of Oi's lending business, Oi Paggo. The test showed the model was an accurate predictor of default—its scores were positively correlated with default across the lending portfolio. The score could be a useful complement to a credit underwriting effort even if it would not replace it. Experian Microanalytics did a similar trial in the Phlippines with MNO SMART and its partner, mBank.

While the initial experience with digital footprints has been primarily with credit, there is growing interest in developing models for other products. Digital footprints could help match people to types

of insurance and help providers tailor premium levels and payment methods to fit people's abilities and needs. Indeed, digital footprints from mobile phone use could yield two basic types of models: (1) predictive models to design financial products and (2) "propensity" models to be used primarily for marketing. For instance, providers could use the propensity models to match savings products with poor people more likely to save based on their mobile phone usage.

Businesses using digital footprints for financial services

As shown in the Figure 2, companies involved in the use of digital footprints can be categorized based on the type of data—financial or nonfinancial—and whether data are used for predictive or propensity models. Because nonfinancial data cover more people and are less explored so far, the opportunity for continued innovation is on the right-hand side of Figure 2.

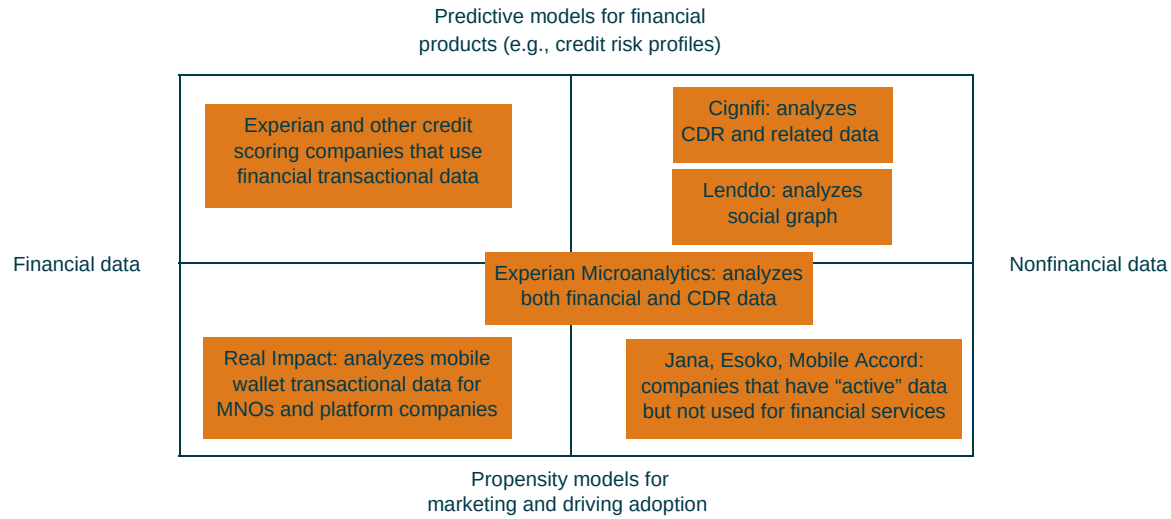
On the bottom left of Figure 2, the opportunity for innovation is in business models. The potential revenue from analysis of financial transactional data supports what are known as "two-sided business models," where the data themselves are considered a bigger source of revenue than direct charges to customers. For example, a "freemium" business model in mobile money would offer basic transactions for free to build transactional volume; once you build sufficient volumes of transactional data, analysis of that data could yield revenues in excess of the revenue foregone by providing the services for free (Kumar and Mino 2011).

Protecting consumers' privacy, security, and ethical use of their digital footprints

Even though there are opportunities for innovation by providers, there remain significant privacy, security, and ethical use concerns with the use of these data. It is also unclear which regulatory body—the telecommunications regulator, banking regulator, or another—has relevant jurisdictional oversight.

Some countries clearly regulate the ownership, use, handling, storage, and transfer of personal data. Regulations in Mexico differentiate between

Figure 2: Example of businesses that use digital footprints for financial services



customers as "data owners" and providers as "data custodians." However, in most countries, data privacy laws are not well developed. It is unclear if poor people even have rights to the digital footprints data generated from their own mobile phone usage and exactly how they would consent to the use of these data. Unless there are adequate protections in place, poor people's data may get commercialized without their consent or knowledge.

Even the world's most sophisticated data companies struggle to protect data they have from criminals looking to exploit that information for financial gain. There are also two sides to the ethical use question. On the one hand, consumers may game the system (e.g., change prepaid airtime purchase patterns) if they are aware of how data are being used. On the other hand, poor people may be denied financial access based on analysis that might treat them unfairly.

Conclusion

The following will need to happen for digital footprints from mobile phone use to make a difference to the unbanked poor:

- Better understanding of the availability and quality of data
- More openness to experimentation by providers, either in partnership with existing analytics companies or through investments in their own home-grown analytics teams
- Improvements in regulation and clear guidance to providers on protection for consumers when it comes to data privacy, security, and ethical use

References

Eagle, Nathan, et al. 2010. "Who's Calling? Demographics of Mobile Phone Use in Rwanda." AAAI Spring Symposium 2010 on Artificial Intelligence for Development (AI-D). http://ai-d.org/pdfs/Blumenstock.pdf

Kumar, Kabir, and Toru Mino. 2011. "Can Mobile Money be Free?" CGAP Technology Blog. http://technology.cgap.org/2011/05/20/can-mobile-money-be-free/

Pakistan Telecom Authority. 2004. Access Promotion Contribution Rules. Pakistan: Pakistan Telecom Authority.

Wired. 2011. "Which Telecoms Store Your Data the Longest? Secret Memo Tells All." http://www.wired.com/threatlevel/2011/09/cellular-customer-data/

AUTHORS:
Kabir Kumar and Kim Muhota

CGAP

July 2012

All CGAP publications are available on the CGAP Web site at www.cgap.org.

CGAP
1818 H Street, NW
MSN P3-300
Washington, DC
20433 USA

Tel: 202-473-9594
Fax: 202-522-3744

Email:
cgap@worldbank.org