

Natural Experiments for Quality of Experience in Networked Systems

Fabián E. Bustamante
Northwestern U.

This short position paper argues that natural experiments and similar alternative experiment designs may offer a promising approach to address the challenges of capturing and experimenting with quality of experience in networked systems.

Networked system research has historically relied on randomized control experiments to draw causal inferences and understand the value of new designs and the effect of interventions. Randomized control experiments, where subjects are assigned at random to treatment or control groups, yield clearly solid inferences since, up to random error, the assignment will balance both groups with respect to all relevant factors other than the treatment and causation.

The scale and criticality of many networked systems, however, make it at best challenging to run randomized controlled experiments at even fractions of these systems' scales and virtually impossible to capture users' assessment of their experiences (with either subjective or objective metrics). In many domains – from broadband services to cell networks – these random experiments are either too expensive or outright impossible to execute for ethical or practical reasons, leaving researchers unable to control the assignment of individuals to conditions or to determine the actual conditions themselves.

This has been a long, well understood problem in a range of fields, from epidemiology and social health to sociology and economics. Observational studies – where the subjects rather than the experimenter assign themselves to the different groups – are a good alternative in some cases. In the context of broadband services again, one could study the experience of users of alternative technologies (e.g., DSL, cable or fiber) by comparing the normalized demand of subscribers of the different technologies. A common problem with observational studies is the impact of confounding factors, factors other than the treatment that make a difference between treatment and control group and affect the response being study. In this example, subscription to a given service and technology could be partially guided by perceived need, age, income or simply availability (i.e., there is no alternative service available). A direct comparison between users without considering, say, income and availability, will be misleading and may suggest correlations when there are none. A typical way to handle this is to make comparisons separately for smaller, more homogenous groups (e.g., households with income above the national average in areas with available alternative services) but this comes at the cost of study group sizes and thus confidence in the results.

Beyond basic observational studies, a commonly use experiment design in such fields is what is generally referred to as “natural experiments”. Natural experiments are a type of observational study where assignments to treatment and control groups is *as if* randomized by nature. They are probably better known in association with John Snow's 1860 study of cholera rates in London's households served by different water companies. On August 1854, the Soho district suffered a major outbreak of cholera that

would eventually take over 600 recorded deaths. John Snow, a London physician, identified the source of the outbreak as the nearest public water pump, which he identified using a map of deaths and illness. In this example, Snow discovered a strong association between sources of water and health problems due to cholera. He found that the water company that supplied water to districts with high death and illnesses rates obtained the water from the Thames downstream from where raw sewage was discharged into the river. By contrast, districts that were supplied water by a different company, which obtained water upstream from the points of sewage discharge, had fewer problems. Given the near-haphazard patchwork development of the water supply in mid-Nineteenth Century London and considering that the exposure to the polluted water was not under the control of any scientist, this exposure has been recognized as being a natural experiment.

I posit that this and other experimental designs may offer a viable alternative to otherwise too expensive or nearly impossible randomized control experiments. As John Snow, by carefully matching users in a treated group with similar users in an untreated group experimenters can manually simulate random or as-good-as-random assignment, ensuring that differences are evenly distributed between the two groups. This allows one to more accurately infer whether or not the relationships observed in a given study are likely to be causal. For example, to test if bandwidth capacity affects user demand, one could pair users that are similar in terms of connection quality and broadband market and check if users with higher capacity generate relatively more traffic.

At Northwestern we have been exploring the use of these and other experiment designs in the context of broadband services and next generation cellular networks. What new inferences are we able to draw relying on such experimental designs? What (other) type of designs can be “imported” to the domain of networked systems? Is there a way to “bind” random controlled and natural experiments in our context? Our recent IMC paper presents an application of these ideas with the first study of broadband services in their broader context, evaluating the impact of service characteristics (such as capacity, latency and loss), their broadband pricing and user demand. A key contribution of that work is a methodology for combining broadband measurement and retail price datasets along with the application of natural experiments to get to a problem otherwise impossible to tackle. In a more recent work, we are exploring the impact of broadband service reliability on users’ QoE partially through the use of these alternative experiment designs.