Stakeholder Network Dynamics Supporting Improved Community Management of Arsenic Filters in West Bengal

Long Paper
Authors Duncan McNicholl, PhD Candidate, University of Cambridge, Department of Engineering, Centre for Sustainable Development, United Kingdom, drm60@cam.ac.uk, +44 754 662 8114
Abstract
This paper presents findings of stakeholder dynamics around community managed arsenic removal facilities in West Bengal, India, to identify conditions that help local management to improve performance. Egocentric network mapping, combined with stakeholder perspectives on important causes, is first used to identify potential drivers of improved performance. These perceptions are then tested in network data to understand their significance. These methods allow triangulation of conditions that are both perceived as important for helping community management performance to improve, and quantitatively visible from a network perspective. Fuzzy Set Qualitative Comparative Analysis is then applied to interpret the relevance of the conditions. Five conditions are identified as relevant: good local relationships; committees engaging and receiving feedback from community; consumers paying a water tariff; a paid caretaker managing facility; and a paid network of delivery agents. These findings are specific to fifteen community networks studied in West Bengal, but the identification of these conditions as relevant dynamics helping community management to improve performance is likely to have broader applicability in other contexts.

Introduction and Aims
Successful community management of rural water supply is about more than technology, or the arrangements of the initial set up. There is a need to understand the ongoing experience of stakeholders interacting, and how these relationships support improved management of services (Stein et al. 2014). Management of community arsenic removal filters in West Bengal, India, offers an appropriate case study (German et al. 2014). Naturally occurring arsenic in aquifers poses a serious challenge in parts of West Bengal where arsenic groundwater concentrations exceeds safe levels. This creates an added challenge in providing safe water access beyond safeguards against pathogens. Community managed arsenic filters are a response to this challenge, and a number of community committees and operators are improving performance, in some cases even 10 or more years after facility installation (Sarkar et al. 2010).

Given that each context contains myriad stakeholders and interactions, it becomes important to understand the full breath of relational influences beyond the scope of any one project. This research was therefore designed to understand stakeholder interactions around community arsenic filters in West Bengal from a network perspective, then to identify dynamics in this network that can explain processes helping performance to improve. The research had two objectives:

1. Identify dynamics in stakeholder networks around community-managed arsenic removal facilities.
2. Identify which dynamics are most important for helping community management to improve performance.

Methods
A network approach was used to understanding stakeholder interactions and dynamics that support performance improvements. This relied on stakeholders individually identifying who they interact with, characterising these relationships, and describing dynamics that they perceive as most important. This relies on the assumption that each stakeholder holds a valid perception of reality, and the aggregation of these can identify stakeholder network dynamics and their significance (Herz et al. 2014).

Data were collected using egocentric network mapping whereby stakeholders independently identify who they directly interact with and how. New stakeholders identified provide a referenced recommendation for subsequent interviews. This logic extends interviews to cover all stakeholders that are both well-connected and influential in the area when basic criteria for inclusion are applied.

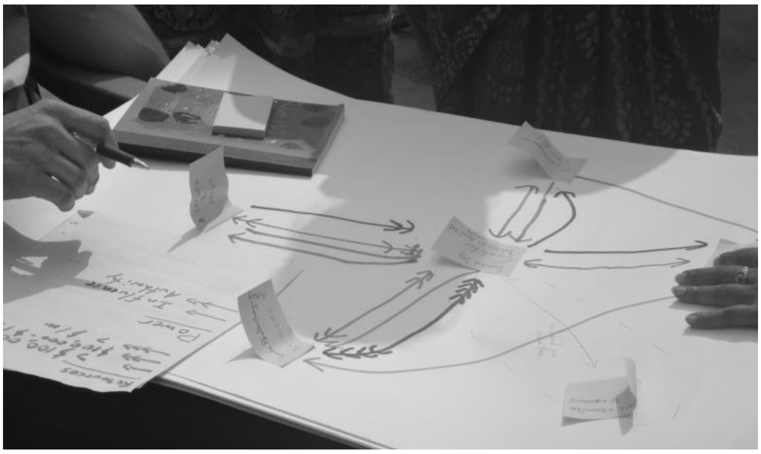

Figure 1 – Egocentric network mapping during an interview

Individual interviews therefore formed the basis of the research methods, and data collected from these were eventually aggregated to produce networks for analysis. Interview participants are first asked about who they interact with on water supply issues. These are written on post-it notes and placed in a circle on a piece of flip chart paper with the stakeholder being interviewed represented in the middle. Participants are then asked to describe their connections to each stakeholders using a framework of defined link types that draws on existing definitions of social power (French & Raven 1959; Raven & French 1958). These have been adapted to familiar terms in the rural water sector that, after pilot testing in multiple countries, have been found to be consistently understood across cultures, contexts, and levels of education. The four categories are: information, skill transfer, resources, and power. Each category is further divided into sub categories designed to represent the relative strength of these connections.

Following drawing of their egocentric network, each participant is asked to comment on relational dynamics that they perceive as helping the management of water services to improve. The question is broad, and is designed to allow the respondent maximum flexibility to describe what they perceive as most important. Participants are encouraged to describe as many factors as they perceive without prompting by potentially leading questions from the researcher. These commentaries are audio recorded for later transcription and analysis.

Sampling
Research locations were recommended by partner organisations intimately familiar with the performance of different community managed arsenic removal facilities. Their recommendations and documentation of performance allowed the selection of a range of performance of community caretakers and committees, ranging from successful and improving, to declining performance, to outright failed. These criteria drew on available data of committee finances, water quality, frequencing of upgrading filter media, and number of local consumers. A ratio of 4:1 successful to less successful cases was desired to emphasise exploration of dynamics that correspond to positive outcomes, while still retaining less successful cases for some comparative analysis. A total of fifteen arsenic removal facilities and their surrounding local networks were included in the study.

Stakeholders at the core of these locations formed the starting point for interviews. These were either the local caretaker or representative of the managing committee. Other starting points included local users, and the organisations interacting with these committees. Further stakeholders were identified for follow up interviews through recommendation during initial interviews. New stakeholders had to be identified as influentially relevant to the local sector, and connected to at least two other stakeholders at the local level, or at least three others at the state level of the local hierarchy. This logic was repeated until all stakeholders identified as relevant and influential were interviewed.

Data Aggregation

Data from the egocentric networks drawn by participants can be aggregated to produce a whole network (Christopoulos et al. 2011). Individual stakeholders become "nodes" in network terminology, and their connections become "edges" (Scott 2013). This includes stakeholders that were identified as involved in the network, even if they were not sufficiently relevant to warrant follow up interviews. These nodes and edges can be input into network software to create a whole network. The open source software Gephi was used in this analysis (Bastian et al. 2009). In some cases stakeholders would mention each other separately when describing a common relationship. Conflicts in these perspectives were resolved by taking the average of two reported edge weights.

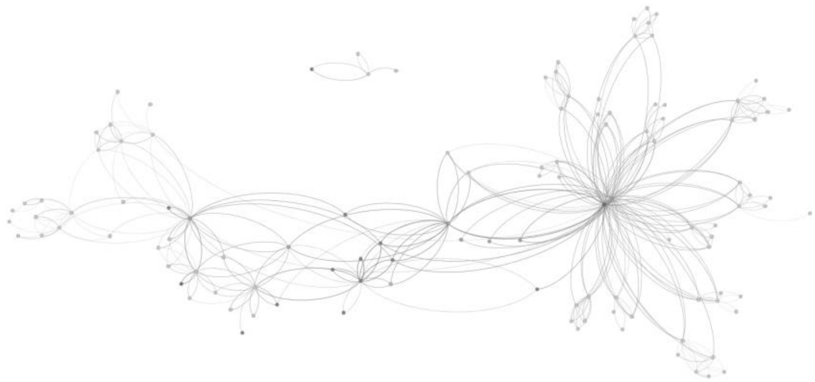

Figure 2 – Whole stakeholder network of aggregated nodes and edges

Transcriptions from verbal descriptions of networks were coded to identify where either specific relationships or dynamics were identified by stakeholders observing that part of the network. This allows the identification of perceived causes of positive outcomes by participants themselves instead of potential causes being defined by the researcher. This identification of factors by interview participants is an important precursor to the use of the Qualitative Comparative Analysis that follows. These two data sets – coded verbal responses and whole network graphs – can be used to triangulate the intersection of multiple perspectives of positive causes and their network characteristics.

Analysis therefore follows a logical sequence. First, perspectives on positive dynamics are sorted to identify where multiple stakeholders perceive the same positive dynamic as causal from different vantage points. Secondly, the description of this dynamic is related by the researcher to a corresponding property to investigate in the network. For example, a description of a "well connected network" could be translated into an assessment of the network diameter, which measures how many stakeholders separate one side of a network from the other – also known as degrees of separation. Finally, drawing on Qualitative Comparative Analysis (QCA), these dynamics can be described as conditions. Each can be a "crisp" set that simply describes a condition as present or absent, or a "fuzzy" set that quantifies a network property in a continuous set (Ragin 2008). Each condition can therefore be described quantitatively from network properties. Similarly, quantitative values can be attributed to the performance of different community management units. These quantitative descriptions allow a mathematical investigation of how network conditions relate to the performance of community management structures. A further step could investigate different configurations of conditions, and how these relate to the desired outcomes (Ragin 2006). The scope of this paper, however, is limited to simply identifying individual factors that are perceived as helping community management structures improve performance, and how these individual conditions relate to observed outcomes.

An important limitation of QCA is the difficulty of identifying which factors are more important than others. This is because QCA looks at configurations of factors that may have codependencies. Most case study stakeholders in this investigation exhibited similar configurations of factors, and the research could be bolstered in the future by studying other stakeholders with similar configurations that are not successful, or different configurations that also lead to success. It should therefore be understood that numerical values in this paper describing different conditions only comment on whether or not a condition is important, but not its relative importance.

Analysis – Step One: Triangulating Perceptions from Verbal Responses
Fragments from interview verbal responses are coded to identify descriptions of positive dynamics, the themes they refer to, and the parts of the network where the observation applies. These codes are then sorted to identify where different stakeholders perceive the same type of positive dynamic from different perspectives. For example, the concept of community water users providing feedback to either local caretakers or committees was a recurring dynamic identified from both the perspectives of users and facility managers:

Local respondent A:

> *"What he is saying is that he arranges a meeting every day, sorry every month, so like, at that meeting he calls upon everybody at the meeting and he listens to the solutions and problems they are facing. Like what are the….how they can do better in the future. So he listens to the problem and if there is a solution so he can do the solution either by spending money or by spending people in the system. So this is how it works."*

Local respondent B:

> *"So he is telling that he gives the water to the consumers and then if the consumers give the feedback to him, then he report it back to the operator."*

These quotes describe a feedback relationship from consumers to the facility managers. Identification of this dynamic from independent perspectives lends causal weight to the relevance of the dynamic.

The number of such overlapping perspectives can be quantified to help identify dynamics that are perceived as the most important. In the case of feedback, N = 6. This same process is applied to all commentary on dynamics interacting at the community stakeholder level. This produces the following group of dynamics at the community level that were perceived as important from multiple perspectives.

A. **Good local relationships** (N = 8). "*See basically you can see our relationship, he is the person who is giving the water to the family, to the local familieis, and he is our very close neighbor. So we are maintaining a very healthy relationship with our consumers. Because you can say everyone is our club member. So in that case we are very close to each other.*"
B. **Committee engaging and receiving feedback from community** (N = 6). "*And the appointed caretaker in most cases they pay an honourary amount so the caretaker from the committee fund.*"
C. **Consumers paying water tariff** (N = 2). "*So this is how the community is getting developed and also getting help by the money the consumers is paying for that water.*"
D. **Paid caretaker managing facility** (N = 2). "…*and with that money he uses that money for giving a salary to [the operator] and [he] is the caretaker and he also caretakes this facility. He gives salary for that.*"
E. **Paid network of delivery agents** (N = 1). "*And kind of carrying charges are being paid by the beneficiaries. So this kind of system really helped.*"

Analysis – Step Two: Relating Perceptions to Network Properties for Investigation
Important dynamics may be observable in multiple local networks. It is therefore helpful to identify corresponding network properties to investigate from the described dynamics across all local networks. Network properties to investigate are defined by the researcher using logical identification of appropriate network parameters. In some cases these are continuous numerical properties, such as network density or betweeness centrality. Others properties can be more straightforward, such as quantifying the strength of a feedback relationship between users and providers. Fortunately, the weighted network graphs can allow quantitative investigation of qualitative dynamics such as feedback, since these weights have been directly reported by respondents in interviews, and are commonly verified from multiple perspectives.

For example, feedback between users and facility managers can be directly filtered from network relationship data to identify the maximum strength of the information relationship between usersand either caretakers or committee members. This is a logical investigation to understand the feedback relationship from a network perspective, and is easily performed on network data. Similarly, other positively described dynamics can be translated into appropriate network parameters to investigate as detailed in Table 1.

Table 1 – Network dyanmics described during interviews and their corresponding network properties to investigate

Described Dynamic	Network Investigation
Good local relationships	Directed network density of local stakeholders
Committee engaging and receiving feedback from community	Weight of strongest information edge between users and either local committee or caretaker
Consumers paying water tariff	Presence of any resource edge from users to either local committee or caretaker
Paid caretaker managing facility	Presence of any local resource edge with the caretaker as the target
Paid network of delivery agents	Presence of any local resource edge with delivery agents as the target

Analysis – Step Three: Network Investigation and Corresponding Outcome
With the relevant parameters defined, the network can be investigated to understand how conditions relate to specific outcomes. For example, the idea of "good relationships" can be translated into an investigation of the proportion of total possible connections that exist in a local network. This is measured as "density", which can either take the direction of the edges into account or not. All references in this paper are to directed density since the direction of these connections matters. Figure 3 presents a local network around one of the arsenic removal sites studied and its corresponding network density. Because four types of edges are included in the network, the greatest possible network density value is 4.0.

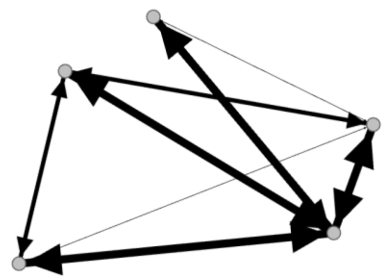

Figure 3 – Directed graph density of a local stakeholder community around an arsenic removal facility

The concept of feedback between users and facility managers can be measured more directly. The weight of information edges between these two network nodes is quantified on a scale of 0 to 3, as reported by stakeholders interviewed. These parameters form the basis of conditions in Fuzzy Set Qualitative Analysis. Each quantitative scale can be normalized on a range from 0 to 1. This means that the raw value of network density can be divided by four, and the raw value of the feedback link can be divided by 3 to make the maximum value 1 and the minimum value 0. Some sets are more straightforward still, such as the presence of a paid caretaker. This is defined as either existing (value = 1) or not (value = 0). This investigation in network data simply requires looking for a resource link leading to a local stakeholder defined as a caretaker.

Quantifying the value of the outcome is the next step. This allows comparison of how quantitative values of different conditions relate to the outcome of improving community management performance. Values for outcomes were assigned on a scale of 0 to 1 for failed and successful units respectively. Some units fell between these extremes. One had been declining in performance and had recently begun to improve; another was still functioning, but declining in performance. These two cases were assigned outcome values of 0.6 and 0.4 respectively to indicate either side of the midway mark between 0 and 1.

With conditions and outcome values quantified, each of these conditions can be similarly investigated to understand two important properties of Qualitative Comparative Analysis: consistency and coverage. Consistency measures how well the presence of a condition coincides with better outcomes. Coverage measures the proportion of successful outcomes that are observed containing a condition. Coverage is important because there can be many pathways to a positive outcome. A high degree of coverage means that most positive outcomes occur with the presence of this condition (Ragin 2006).

These parameters are distinct from quantitative regression analysis because they investigate the extent to which conditions or outcomes are contained within each other instead of quantitative adherence to a curve. This is helpful to describe conditions that are necessary for a desired outcome, but insufficient themselves. Conversely, they can describe conditions that can help to achieve outcomes, but are not strictly necessary because other pathways to success exist.

Main results and lessons learnt

To summarise, each condition was first identified through participant perspectives. These were filtered to identify where multiple perspectives agreed, forming the basis for what to investigate in Qualitative Comparative Analysis. These conditions were then related to a quantitative property that could be investigated in network data. Each condition was then assessed against the outcome of community management performance to quantify how well the presence of a condition explains the presence of the desired outcome. Consistency scores above 0.75 are typically considered significant (Ragin 2006). Coverage scores, on the other hand, explain how necessary a condition is to achieving an outcome. This means that the presence of a condition might help create an outcome and therefore have a high consistency score, while also having a low coverage score that indicates other pathways to the outcome exist.

Table 2 presents an example of this analysis showing how outcome values and directed network density values are observed in local networks at different locations. Since the raw network density maximum value is 4, dividing by this figure calibrates network density values into a fuzzy set from 0 to 1. These values are then used to calculate coverage and consistency for the condition "good local relationships."

Table 2 – Calculating consistency and coverage for directed network density

Location	Fuzzy Outcome Value	Network Density Value	Fuzzy Network Density
A	1	1.75	0.4375
B	1	0.892	0.223
C	1	1.2	0.3
D	1	1.7	0.425
E	1	1	0.25
F	1	1.45	0.3625
G	1	1.032	0.258
H	1	1.45	0.3625
I	1	0.917	0.22925
J	1	1.45	0.3625
K	1	1.19	0.2975
L	0.4	1.167	0.29175
M	0.2	0.75	0.1875
N	0.6	1.834	0.4585
O	0	0.584	0.146
CONSISTENCY			**0.97**
COVERAGE			**0.36**

The consistency score of 0.97 found for the property of network density indicates that the presence of higher network density values strongly coincides with improved outcomes. The relatively low coverage score, however, indicates that network density, although helpful, may not be completely necessary to achieving improved outcomes.

Similar calculations are repeated for all five conditions. Table 3 presents the summary of findings. All five conditions have consistency values well above the 0.75 threshhold meaning that the presence of each condition is a good predictor of a favourable outcome. This finding is not symmetrical, however. A low coverage score, notably for the network density condition used to describe "good local relationships", indicates that the absence of the condition does not necessarily undermine the outcome. There can simply be other recipes for success.

Table 3 – Summary of consistency and coverage values for the five conditions

Condition	Consistency	Coverage
Good Local Relationships	0.97	0.36
Committee engaging and receiving feedback from community	0.92	0.89
Consumers paying water tariff	0.92	0.98
Paid caretaker managing facility	0.87	1.00
Paid network of delivery agents	1.00	0.82

Conclusions

This research sought to identify network dynamics around instances of improved performance in the management of rural arsenic facilities in West Bengal, and interpret the relative significance of these dynamics. Five conditions were identified as potentially important causes, and evaluation of the relationship between their presence and the desired outcome finds that all five are portential indicators of improving community management performance. Each may have causal significance as well, but isolating these is difficult because the conditions were typically observed together instead of individually. The five conditions are:

1. Good local relationships
2. Committee engaging and receiving feedback from community
3. Consumers paying water tariff
4. Paid caretaker managing facility
5. Paid network of delivery agents

Each condition can be logically related to helping the management of community arsenic filters improve as a potential cause. Good relations in tightly knit networks can suggest dynamics such as accountability, trust, and feedback that may help performance to improve. Similarly, having dedicated facility managers and delivery agents that are paid by users can further help to strengthen roles, responsibilities, and accountability in these local networks. Failure in cases lacking many of these conditions is easy to imagine. The absence of feedback from users, the absence of a paid caretaker, and the lack of payment for water, all seem logical conditions that would undermine success. This assumption is supported, although not empirically tested with the same methodology, by the failure of other arsenic removal facilities in the region that failed quickly in the absence of formalized management structures (Hossain et al. 2005).

These findings can be used to help identify the presence of conditions that support the improvement of community management in this context, and potentially to help introduce conditions found to be absent in order to improve the sustainability and performance of such sites. The identification of these conditions, and the confirmation of their significance, is a useful basis for further investigation because these are informed directly by stakeholder experiences and not *a priori* judgments of the researcher. The identification of these conditions, and their relevance towards successful configurations described in this paper, can hopefully help both researchers and practitioners to better understand and support communities managing safe water provision locally.

Acknowledgements

This research was made possible by Professor SenGupta of Lehigh University and the Society for Technology with a Human Face organisation, as well as the Professor Gupta of IIEST, Shibpur, and the SATHEE organisation. Their involvement was critical for identifying suitable case study locations and providing data on the performance of community management *units. The research also benefitted greatly from the expert guidance of Dr. McRobie and Dr. Cruickshank of Cambridge University.*

References

Bastian, M., Heymann, S. & Jacomy, M., 2009. Gephi: an open source software for exploring and manipulating networks. In *International AAAI Conference on Weblogs and Social Media.*

Christopoulos, D. et al., 2011. Connections Journal 2011. *Connections*, 31(1), pp.1–52.

French, J.R.P. & Raven, B., 1959. Bases of Social Power. *Control*, pp.385–385.

German, M., Seingheng, H. & SenGupta, A.K., 2014. Mitigating arsenic crisis in the developing world: Role of robust, reusable and selective hybrid anion exchanger (HAIX). *Science of the Total Environment*, 488-489, pp.547–553. Available at: http://dx.doi.org/10.1016/j.scitotenv.2013.10.092.

Herz, A., Peters, L. & Truschkat, I., 2014. How to Do Qualitative Structural Analysis: the Qualitative Interpretation of Network Maps and Narrative Interviews. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 16(1), p.no pages. Available at: http://www.qualitative-research.net/index.php/fqs/article/view/2092.

Hossain, M. et al., 2005. Ineffectiveness and Poor Reliability of Arsenic Removal Plants in West Bengal, India. *Environmental Science & Technology*, pp.4300–4306.

Ragin, C.C., 2008. Qualitative Comparative Analysis Using Fuzzy Sets (fsQCA). *Configurational Comparative Analysis*, pp.87–121.

Ragin, C.C., 2006. Set relations in social research: Evaluating their consistency and coverage. *Political Analysis*, 14(3), pp.291–310.

Raven, B.H. & French, J.R., 1958. Legitimate power, coercive power, and observability in social influence. *Sociometry*, 21(2), pp.83–97. Available at: http://www.jstor.org/stable/10.2307/2785895.

Sarkar, S. et al., 2010. Evolution of community-based arsenic removal systems in remote villages in West Bengal, India: Assessment of decade-long operation. *Water Research*, 44(19), pp.5813–5822. Available at: http://dx.doi.org/10.1016/j.watres.2010.07.072.

Scott, J., 2013. *Social Network Analysis*, London: SAGE Publications Ltd.

Stein, C. et al., 2014. *Advancing the Water-energy-food Nexus : Social Networks and Institutional Interplay in the Blue Nile*, Colombo, Sri Lanka. Available at: https://www.sei-international.org/publications?pid=2573.

Contact Details

Name of Lead Author: Duncan McNicholl
Email: drm60@cam.ac.uk