[Supplementary material]

Networked medieval strongholds in Garhwal Himalaya, India

Nagendra Singh Rawat[1,*], Tom Brughmans[2], Vinod Nautiyal[1] & Devidutt Chauniyal[3]

[1] Department of History and Archaeology Hemvati Nandan Bahuguna Garhwal (Central) University, India

[2] Centre for Urban Network Evolutions (UrbNet), Aarhus University, Denmark

[3] Department of Geography, Hemvati Nandan Bahuguna Garhwal (Central) University, India

[*] Author for correspondence: ✉ nagendra.rawat957@gmail.com

Table S1. List of all 193 strongholds in Garhwal Himalaya identified by Rawat (2017), with location, the sources through which they were identified, geological and topographic information about their locations, and results of 15km and 25km network degree, betweenness, closeness and Louvain modularity analyses (see the separate Excel file).

Method line-of-sight network creation

The Digital Elevation Model (DEM) of the study area used was procured through the region's *Uttarakhand Space Application Centre* (U-SAC), Dehradun (30m horizontal resolution). Visibility networks were calculated using the open source *Viewshed Analysis* plugin v0.5.4 by Zoran Čučković (2016) in QGIS v2.18.2. Earth curvature was taken into account (atmospheric refraction set to 0.13).

We created these line-of-sight networks with a range of maximum viewing distances representing different hypotheses. A 3km maximum viewing distance is used to reflect the distance at which individual humans can be discerned; a 15km viewing distance is used to reflect the distance at which fire and smoke signals are definitely still visible (Rego & Catry 2006); 25km is used to reflect the maximum distance at which large fire and smoke signals are visible (Rego & Catry 2006); 50km is used to reflect the visibility of huge fire, light or smoke signals under particularly exceptional circumstances.

For the vast majority of forts and fortalices no information is available about the vertical elevation of structures, let alone the height of the signalling platform. However, all fortalices and forts are very large structures (Figures 1 & 2) and can be assumed to rise above the surrounding vegetation. In this study we assume an *observer*'s eye level is at an elevation of 10m above the topography elevation, and the elevation of the *observed* signal (flags, smoke column, fire) also at 10m. We selected these values because they reflect our assumption that signals and observers would be located on top of large fortress or watchtower structures above the canopy of surrounding vegetation. Variation in the height of the observed flag, smoke or fire signals can be expected but experiments performed with signal heights of 5m and 12m only led to very minor changes in the resulting visibility networks. Moreover, the use of 10m elevation for both observer and signal has the added advantage that we can assume intervisibility of all lines of sight (Conolly & Lake 2006: 229–30). The resulting networks are therefore undirected: nodes represent site locations (both as observation location and as observed location) and edges represent bi-directional lines-of-sight. We performed line-of-sight analyses from and to the highest raster cell within a 90m radius from site point locations, representing the unknown extent of the site features and the assumption that an observation point would be located on the locally highest point.

This method results in four line-of-sight networks with maximum viewing distances of respectively 3km, 15km, 25km and 50km.

The 3km network is extremely fragmented: very few site pairs are within 3km of each other. This means that the theorized visual signalling network could certainly not have been established to observe individual humans or signalling using the body. On the other side of the spectrum, the 50km network is very dense and well connected, but at such vast distances flags would certainly not be visible, and only huge fires and smoke columns would be visible (cf. forest fires). For these reasons we do not interpret the 3km and 50km networks in any detain in this paper (although their network metrics are still presented in Table 1 as a reference). The maximum viewing distances of 15km and 25km best represent a functional visual signalling system and we will therefore focus in particular on the analyses of these networks in this paper.

Definition network statistics

A range of network analysis techniques has been applied to these networks (Figure 4). The results are presented in Table 1 and Figures 5–7. Here we define all metrics applied.

Nodes: the number of nodes (or entities) in the network. In our study, nodes are used to represent sites (both forts and fortalices).

Edges: the number of edges (or lines) in the network. In our study, undirected edges are used to represent lines-of-sight between observers at pairs of sites.

Density: the fraction of the number of edges that are present in the visibility network to the maximum possible number of edges.

Average degree: a node's degree is the number of edges it has (i.e. the number of lines-of-sight related to a stronghold). The average degree of a network is the sum of all nodes' degrees divided by the number of nodes.

Isolated nodes: nodes that are not connected to other nodes via edges are isolated (i.e. sites with no lines-of-sight related to them).

Isolated forts: the number of forts (and not fortalices) that are not connected to other sites with lines-of-sight.

Components > 2 nodes: a component is a subset of the visibility network where all nodes in the subset can be connected to each other via paths, but none of the nodes in the subset can be connected to other nodes outside the subset. In this study we count the number of components that consist of more than 2 nodes, to give a representation of the number of groups of nodes that could have functioned as a signalling network (assuming a sensible signalling network should consist of at least three nodes such that information can be passed from between a pair of nodes via an intermediary).

Size largest component: The number of nodes in the largest component.

Degree distribution: The frequency distribution of the degrees of all nodes in the network (i.e. how many nodes have degree 1?, etc.).

Betweenness centrality distribution: the frequency distribution of the betweenness centrality of all nodes in the network. A node's betweenness centrality is the fraction of the number of shortest paths passing through this node over the number of shortest paths between all node pairs in the network. It is used to determine how important a node is in mediating information through the network, as a go-between.

Louvain modularity community detection: modularity refers to the identification of groups of nodes (so-called communities) where the density of nodes within the group is significantly higher than the density we would expect given a random network creation process. Louvain modularity is a particularly commonly applied modularity method that aims to maximize modularity scores for communities (Blondel *et al*. 2008). It results in the identification of communities in the network (Figure 7).

Results network statistics

Networks were analysed using Visone v2.17 (Visone 2016). Table 1 presents summary network statistics for visibility networks with maximum viewing distances of 3km, 15km, 25km, and 50km (the 15km and 25km networks are shown in Figure 4). The visibility networks have a very low density, which is to be expected for geographical networks where the nodes are dispersed over a huge mountainous area. The network with a maximum viewing distance of 3km is particularly sparse with a high number of isolated nodes: very few forts and fortalices are located such that individual humans would be visible between them. However, the visibility networks become denser with far less isolates at distances of 15km and larger: most forts and fortalices are located such that visual signals are visible from at least one other stronghold. Isolated nodes are more often fortalices than forts in absolute numbers, although the proportion of isolated forts over all isolates is similar to the proportion of the total number of forts over all sites (36/193). This suggests forts are not more commonly included in the visibility network than fortalices.

At a maximum viewing distance of 3km the visibility network is largely disconnected and only five small components consist of more than 2 sites. This suggests that no sites in the study area seem particularly well located to observe individuals at other sites, or to function as a signalling network over particularly short distances.

The visibility networks with a maximum viewing distance of 15km and 25km are most likely to reflect the structure of a possible signalling network since at these distances fire and smoke signals are still visible whilst allowing for observation points to be spaced over a large area. The largest component of the 15km network includes 79 sites (41 per cent of all sites). The 25km network is far denser with no less than 120 sites (62 per cent of all sites) included in the largest component: information can be shared from any of the 120 sites to any other. Table 2 (columns marked “Observed”) further shows that both the 15km and the 25km networks have very high counts for all network configurations discussed in the method section and representing our theories of a functional visual signalling network.

We can conclude that both the 15 and 25km networks show the network structural features one would expect to see in a visual signalling network.

Results degree and betweenness distribution

What was the role of particular sites in this hypothetical signalling network? In our study we have used a range of statistics describing roles of visual control (represented using node degree) and mediating visual communication (represented using node betweenness centrality) of all nodes in the visibility networks. We explored the frequency distributions for these two metrics, to explore whether there was a typical average value or whether there groups of nodes with very different roles. Our results support the latter: a small number of nodes had a very high degree or betweenness centrality and the vast majority of nodes have very low scores. This means that not all strongholds served the same purpose in our theorized visual signalling network. Some strongholds could effectively disseminate information immediately to a high number of others due to having many lines-of-sight connected to them (i.e. a high degree) (as also shown by the significantly high counts of the *Alternating-star* configuration; Table 2). Some strongholds could control the flow of information throughout the signalling system effectively by acting as crucial intermediaries, often bridging different parts of the network (i.e. a high betweenness).

Figure 5 presents the degree distributions of the visibility networks with maximum viewing distances up to 15km and 25km. The overall degree distributions for both networks is skewed to the right, indicating that from the vast majority of sites a very low number of other sites can be seen whereas from a few sites a much higher number of other sites can be seen. Crucially, there is little difference between the degree distributions of forts and fortalices, suggesting that both forts and fortalices could have played a role as hubs in a possible communication network and could have visually controlled surrounding sites. Figure 6 presents the distribution of betweenness centrality scores, which illustrates that the sites with high scores can be both forts and fortalices, but that fortalices are always the most important intermediaries.

Method and results community structure

To further explore our second research question, concerning whether the visibility networks could help reveal sets of strongholds that might have formed a local chiefdom or mandala, we apply the Louvain modularity technique (Blondel *et al*. 2008). This allows us to identify those sets of strongholds as so-called communities, which are more intervisible with each other than with any other strongholds (i.e. subnetworks with a particularly high density).

The results for the 15km network reveal communities on the borders of the study area that could have functioned as short distance visual signalling clusters, effectively communicating possible enemy incursions locally (Figure 7a). However, these communities become integrated with those at the centre of the study region when the method is applied to the 25km network (Figure 7b). The identified communities do not correlate with watersheds or rock formations (Table S1). The role of major forts in these communities is explored in the discussion section of the paper.

Definition and method network configurations

If these 15km and 25km networks functioned as visual signalling systems, then they would consist of the patterns that enable signals to be spread effectively. A new method for evaluating the presence of such patterns has been used in our research (Brughmans *et al.* 2014; Brughmans & Brandes 2017). This method allows us to define the patterns we expect to find in an efficient visual signalling network according to our theory, these patterns are represented as so-called configurations (Figure 8), to count the frequency of these configurations and to evaluate whether or not they are common. If these theorised patterns are particularly high in our 15km and 25km networks, then this adds support to our theory that they functioned as a visual signalling network, because they include the patterns we would expect to see.

We expect to see the following configurations in visual signalling networks:

Isolates (Figure 8a): a visual signalling network would ideally include all strongholds in the region, so we expect to see few isolated nodes.

3-paths (Figure 8d): to enable the sharing of information along chains of intervisible strongholds, we expect to see paths of at least three or more strongholds. This reflects the ability to share information between a pair of strongholds via intermediaries, a key characteristic of a visual signalling network.

Alternating 2-paths (Figure 8b): ideally, a visual signalling network would not lose its primary function if one stronghold is removed (for example because it is attacked or because it has technical difficulties preventing it from signalling). For this reason, we expect robust visual signalling networks to have a high number of alternative paths connecting pairs of nodes.

Alternating-star (Figure 8c): we expect some strongholds to be better connected to others by being able to share information directly to many other strongholds. The *alternating-star* configuration represents this phenomenon and a tendency for the degree distribution to be skewed.

In light of our theory about the role of major forts in this visual signalling network and our fourth research question, we additionally explore whether the *forts* tend to be engaged in specific configurations:

Fort activity (Figure 8e): we expect forts to be included in the network rather than being isolated.

Fort 2-path O1au (Figure 8f): forts might have a tendency to act as intermediaries in visual signalling.

Fort triangle-1U (Figure 8g): alternatively, forts might be part of dense local clusters and occupy a similar structural position as other neighbouring strongholds, as reflected by a triangle pattern.

Results:

Table 2 (columns marked “Observed”) shows that both the 15km and the 25km network have very high counts for all network configurations. But what does this mean? Does this support our theory of a visual signalling network?

Method and results comparison with random graphs

Are the “observed” counts of the network structures of the 15km and 25km networks high enough to argue they are important and exceptional features of the network? Do these results allow us to argue that these networks were *particularly* well-suited for visual signalling? This will be explored by comparing these results to those from networks with the same number of nodes, edges and density created through a random process, i.e. Bernoulli random graph models. For example, in the case of the 15km network, this means that we generated 1000 networks with 193 nodes connected by 223 edges (i.e. the same as the number of forts and lines-of-sight in the observed 15km network). Crucially, to incorporate an element of the geographical limitations of these sites embedded in a real physical landscape, in a second pair of models we only allow lines-of-sight to be created between pairs of sites within 15km and 25km respectively (using the method based on *structural zeroes* proposed by Brughmans & Brandes 2017). In this part of the method we will therefore statistically compare the ‘observed’ counts of configurations in the 15km and 25km networks, with simulated counts in random networks with the same density, and random networks restricted to lines-of-sight shorter than 15km and 25km respectively (Table 2).

Bernoulli random graph simulations were performed with PNET (Wang *et al*. 2009). We iterated the Bernoulli random graph models 1000 times and counted the resulting numbers of simulated configurations which are all normally distributed (a requirement for performing the significance test we use). Table 2 presents network statistics of such randomly generated networks. We report the mean and standard deviation of simulated configuration counts, and consider these significantly different from the observed counts if the following t-statistic is higher than 2 (an approach recommended by Harrigan 2007: 25–26):

$$t = \frac{(observed\ count - mean\ simulated\ count)}{standard\ deviation\ simulated\ count}$$

The results in Table 2 show that all simulated configuration counts are significantly different from the observed networks (compare 'observed' columns with 'random' and 'random + structural zeroes' columns). The configurations representing our theories about a visual signalling network (*3-path*, *alternating-2-path*, *alternating-star*, *fort 2-path O1au*, *fort activity*) are all much more common in the observed network than in the simulated networks, adding further support to the possibility of this system of forts to have functioned as such. Chains of intervisible sites are very common in the observed visibility networks (*3-paths*) and would have enabled visual signalling over large areas through a sequence of signalling actions. Moreover, we notice this visual signalling network could be robust given the high count of *alternating-2-paths*: a tendency for multiple paths to exist to connect pairs of nodes, such that alternative signalling paths can be used in case one path is compromised to a targeted attack or signalling issue. However, both observed networks, and in particular the 25km network, also have much higher numbers of isolated nodes than one would expect in a randomly generated network. This suggests that a higher number of sites than expected by chance are not part of the hypothesised visual signalling network: a possible past visual signalling network might therefore not have covered the entire study area but rather particularly dense parts, the area of the largest component of these networks. Finally, the forts appear in the theorized configurations particularly frequently, supporting our theory that they tend to be included in the network (*activity*) and played a role as mediators in visual signalling (*2-path O1au*). However, they also tend to be clustered with other sites (*triangle-1U*), which argues against them being the focal points of local networks. They seem to have served similar roles as all other strongholds.

Sensitivity analysis

The network science results presented in this paper assume that all strongholds were contemporary and could have functioned together as a visual signalling network. This assumption is problematic in light of the almost complete absence of dating evidence for any of the forts. Indeed, the results of network metrics are highly sensitive to changes in the number of nodes and edges. We therefore need to understand this sensitivity better: exactly how much can we expect our published network science results to change if we assume just 90 per cent of forts were contemporary, or 80 per cent, or 70 per cent, etc? For example, assuming only 90 per cent of sites were contemporary means 10 per cent of sites and their related edges need to be removed from the network (i.e. we subsample the network). Once these are removed we can calculate our network metrics again and compare them with the original results: did the ranking of nodes according to the network metric change much? If yes, then this metric's results are very sensitive to the missing information about site contemporaneity and we should be very careful to interpret it. If no, then this metric's results are highly robust despite the missing information about site contemporaneity and we can attach more confidence to our interpretations. Doing this procedure just once makes the results highly sensitive to the precise 10 per cent of nodes selected for removal, which is why the procedure needs to be repeated thousands of times to identify the distribution of the change of network results. The final sensitivity analysis results show this distribution of network metric results from thousands of random subsamples as compared to the original network metric results of the complete network.

In this paper, we performed such a statistical sensitivity analysis to evaluate how robust the degree and betweenness results are to missing information about site contemporaneity, by subsampling at 10 per cent intervals between 10 and 90 per cent of all sites. We use the method by Costenbader and Valente (2003) as implemented in R by Peeples (2017). We performed 10000 iterations of the random sub-sampling, and rank-order correlations between network results of the original network and those of the random subsample are calculated with Spearman's ρ.

The results in Figure 9 show that the degree scores are rather robust to removing strongholds, especially in the 25km network: 70 per cent or more of all nodes need to be removed in order to generate huge changes. Betweenness centrality shows more dramatic variation at very high levels of stronghold removal (this is in line with observations that betweenness centrality is typically more sensitive to sub-sampling; Peeples 2017) and should therefore be interpreted with more caution: we see huge changes when 50 per cent of nodes are removed in the case of the 15km network and 60 per cent of nodes for the 25km network. We can conclude that the degree and betweenness results presented in this paper are very robust if 50 per cent or more of strongholds are contemporary, that the degree results are robust with only 30 per cent of contemporary nodes, and that the results for the 25km network are more robust than those of the 15km network. We consider this an acceptable assumption in light of the archaeological and historical context of the fortification phenomenon in Garhwal Himalaya, but new dating evidence collected in future archaeological research will be essential in obtaining more precise network science results.

References

BLONDEL, V.D., J.L. GUILLAUME, R. LAMBIOTTE & E. LEFEBVRE. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*: 1–12. https://doi.org/10.1088/1742-5468/2008/10/P10008

BRUGHMANS, T. & U. BRANDES. 2017. Visibility network patterns and methods for studying visual relational phenomena in archaeology. *Frontiers in Digital Humanities: Digital Archaeology* 4. https://doi.org/doi.org/10.3389/fdigh.2017.00017

BRUGHMANS, T., S. KEAY & G. EARL. 2014. Introducing exponential random graph models for visibility networks. *Journal of Archaeological Science* 49: 442–54. https://doi.org/10.1016/j.jas.2014.05.027

CONOLLY, J. & M. LAKE. 2006. *Geographical information systems in archaeology*. Cambridge & New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511807459

COSTENBADER, E. & T.W. VALENTE. 2003. The stability of centrality measures when networks are sampled. *Social Networks* 25: 283–307. https://doi.org/10.1016/S0378-8733(03)00012-1

ČUČKOVIĆ, Z. 2016. Advanced viewshed analysis: a Quantum GIS plug-in for the analysis of visual landscapes. *The Journal of Open Source Software* 1. https://doi.org/http://dx.doi.org/10.21105/joss.00032.

HARRIGAN, N. 2007. PNet for dummies: an introduction to estimating exponential random graph (p*) models with PNet. Available at: http://ink.library.smu.edu.sg/soss_research/810/ (accessed 05 January 2021).

PEEPLES, M.A. 2017. Network science and statistical techniques for dealing with uncertainties in archaeological datasets. Available at: http://www.mattpeeples.net/netstats.html (accessed 05 January 2021).

REGO, F.C. & F.X. CATRY. 2006. Modelling the effects of distance on the probability of fire detection from lookouts. *International Journal of Wildland Fire* 15: 197. https://doi.org/10.1071/WF04016

VISONE, P.T. 2016. Visone v.2.16. Available at: http://visone.info (accessed 05 January 2021).

WANG, P., G. ROBINS & P. PATTISON. 2009. *Pnet: program for the simulation and estimation of exponential random graph (p*) models. User manual*. Available at: https://static1.squarespace.com/static/57a1436215d5dbbcd2031828/t/581ac45ce58c62432be83a02/1478149214983/PNetManual.pdf (accessed 05 January 2021).