



Backbone Extraction in Archaeological Similarity Networks: Methodology and a Ceramic Case Study from the Amuq Valley

DARIO NINCHERI

MARCO MODERATO

*Author affiliations can be found in the back matter of this article

RESEARCH ARTICLE

ubiquity press

ABSTRACT

This study examines backbone extraction not simply as a filtering technique, but as a diagnostic tool to assess the interpretability of archaeological similarity networks. Using a ceramic dataset from the Amuq Valley during the Bronze Age, we evaluate two backbone extraction methods – the Global Statistical Significance (GloSS) filter and the Disparity Filter – in the context of worst-case conditions: limited typological diversity, cumulative survey data, and extensive chronological overlap. In doing so, the paper introduces backbone extraction to archaeological network research as a formal test of relational selectivity, and asks under what data conditions such selectivity can be statistically recovered.

Application of the Global Statistical Significance (GloSS) filter produces a null result, indicating that the dense similarity structure is statistically indistinguishable from a randomised null model. The Disparity Filter likewise retains no statistically significant relations: the only edge that survives the procedure links the two sites sharing a rare ceramic class (White Slipped Ware), but it arises from a pair of degree-one nodes — a configuration outside the filter's domain of applicability rather than a selective tie. Both global and local filters therefore converge on an empty backbone. This convergence is itself informative: the GloSS filter asks whether edges are unusual *relative to the entire network*, while the Disparity Filter asks whether they are unusual *relative to a node's local neighbourhood*, neither recovers selective structure once the redundancy of the underlying data is taken into account.

Rather than constituting methodological failure, such negative or minimal results offer the empirical boundaries of relational inference from similarity data. We argue that backbone extraction should be understood as a diagnostic framework that exposes the limits of relational inference, rather than a universally applicable method for network simplification. The approach offers archaeologists a principled safeguard against over-interpreting dense similarity networks, and a transferable criterion for distinguishing genuine relational selectivity from background cultural diffusion.

CORRESPONDING AUTHOR:

Marco Moderato

Universitat Pompeu Fabra, ES
marco.moderato@upf.edu

KEYWORDS:

Similarity networks; Backbone extraction; Ceramic analysis; Bipartite projections; Null models; Bronze Age Levant

TO CITE THIS ARTICLE:

Nincheri, D and Moderato, M. 2026. Backbone Extraction in Archaeological Similarity Networks: Methodology and a Ceramic Case Study from the Amuq Valley. *Journal of Computer Applications in Archaeology*, 9(1): 312–328. DOI: <https://doi.org/10.5334/jcaa.281>

1. INTRODUCTION

Over the past two decades, network analysis has become increasingly central to archaeological research. It offers a robust theoretical and methodological toolkit to move from intuitive descriptions of connections to a formal, replicable, and quantitative study of relational systems (Brughmans et al. 2024; Brughmans and Peeples 2017; Collar et al. 2015; Donnellan 2020). Among these approaches, similarity networks—linking sites or contexts on the basis of shared artefact categories—are widely used to explore relational patterns in survey and excavation datasets (Moreno-Navarro 2024; Blair 2024 (in Handbook); Peeples and Mills 2018; Mills 2016; Östborn and Gerding 2014).

Yet archaeological similarity networks are almost invariably dense: when constructed from bipartite projections of site-by-artefact-category matrices, they tend toward near-complete connectivity because most sites share at least some widely diffused categories, compressing weight variability and obscuring selective relational structure (Brughmans and Peeples 2023; Moreno-Navarro 2024; Collar et al. 2015). Despite the prevalence of this problem, archaeology lacks widely adopted formal procedures for testing whether density encodes meaningful structure or merely background diffusion.

In network science, backbone extraction technique has been developed as a formal statistical procedure for reducing dense, noisy weighted networks to sparse subgraphs that preserve only the most significant relational information (Coscia and Neffke 2017; Neal 2014; Serrano et al. 2009). The underlying problem is well documented: many empirical networks — including those derived from aggregated observational data — are dense because most nodes share at least some connectivity, and without filtering, structural analysis tools such as community detection may fail to distinguish meaningful subgroups from background noise (Coscia and Neffke 2017). The aim of backbone extraction is to separate connections that carry genuine signal from those that are consistent with the expected behaviour of a randomly wired network of the same density: each edge is evaluated against a null model, and only those edges whose weights are statistically improbable under that model are retained (Serrano et al. 2009; Neal 2022). The result is a sparser network — the backbone — in which retained edges represent statistically exceptional relationships rather than artefacts of general connectivity. Although widely applied in fields such as economics and social network analysis, these methods remain rarely employed in archaeological research, and their inferential implications for material culture data have not been systematically examined.

This study addresses this gap by evaluating two backbone extraction methods using ceramic assemblage

data from the Amuq Valley Regional Project (Turkey) during the Bronze Age.

Rather than reconstructing interaction networks, we ask three related questions. First, can backbone extraction recover statistically selective relational ties from a similarity network built on survey-derived, binary ceramic data? Second, how do global and local null models differ in what they retain under these conditions? Third, what does a null or minimal backbone reveal about the empirical limits of inference from dense, low-diversity similarity data of the kind routinely produced by archaeological survey? In doing so, we reframe backbone extraction not as a reductive tool, but as a formal test of relational selectivity in archaeology.

2. MATERIALS

2.1 THE AVRP DATASET

The empirical basis of this study is provided by ceramic data collected within the Amuq Valley Regional Projects (AVRP), a large-scale archaeological survey conducted between 1995 and 2002 under the direction of Kutlu Aslıhan Yener (Yener 2005). The project builds on a longer tradition of regional investigation in the Amuq Valley initiated by Braidwood in the 1930s (Braidwood 1937) and developed through subsequent research on the region, including the work of Murat Akar (Akar and Kara 2018). The project systematically documented hundreds of sites across the Amuq Plain and the lower Oronte, generating one of the most comprehensive regional datasets for the northern Levant. The regional survey strategy and its interpretative framework have been extensively discussed in previous publications (Casana et al. 2005; Casana & Wilkinson 2005), which outline settlement distribution, landscape dynamics, and long-term occupation patterns in the Amuq Valley.

Within this framework, Bronze Age ceramics—particularly diagnostic classes such as Red-Black Burnished Ware, Plain Simple Ware, and Slipped Ware—represent one of the main indicators for investigating relationships between sites, as they reflect shared technological practices, distribution circuits, and regionally documented chronological affinities. In the regional sequence, the Bronze Age corresponds to Amuq Phases H–M (Yener 2005; Casana 2009), spanning the third and second millennia BC. Red-Black Burnished (Khirbet Kerak) Ware is diagnostic of the mid-third millennium (Amuq H–I), while the second-millennium sequence comprises the Middle Bronze Age (ca. 2000–1550 BC) and the Late Bronze Age (Amuq M, ca. 1550–1200 BC), the latter recognisable in survey assemblages almost exclusively through rare imported or imitation Mycenaean, Cypriot, and Nuzi wares (Casana 2009). Red-Black Burnished Ware and Plain Simple Ware are the principal classes recorded as chronologically diagnostic for this interval in the survey record.

For the purposes of the present analysis, 35 Bronze Age sites were selected based on the availability of diagnostic ceramic classes (Figure 1).

The four ceramic classes used here — Red-Black Burnished Ware, Plain Simple Ware, White Slipped Ware, and Slipped Ware — are not an analytical reduction adopted for convenience, but the classes that the AVRPPotteryAssessments themselves record as chronologically diagnostic for the Bronze Age (Yener 2005). The analysis therefore inherits the typological resolution of its source rather than imposing one. We acknowledge that this is a coarse-grained representation: ware is only one axis of ceramic variation, and two of the four classes occur at very few sites. This coarseness is not incidental to the study but central to it, since our aim is to test whether relational selectivity can be recovered from precisely the kind of low-resolution, survey-derived categorical data that archaeologists routinely work with; we return to its consequences in Section 5.6.

The dataset used in this study is constructed exclusively from the Pottery Assessments section of the Amuq Valley Regional Project (AVRP) gazetteer. These assessments summarize ceramic observations made during regional survey and provide preliminary chronological attributions for surveyed sites. As a result, several major excavated sites of the Amuq Valley—including Tell Atchana (Alalakh), Tell Tayinat, Tell Judaidah, and Çatal Höyük (Amuq)—are not included

in the dataset, since their ceramic assemblages are not explicitly listed in this survey-oriented section. The network therefore represents relationships among sites documented through the regional survey record rather than the full set of excavated settlements.

The absence of these major sites does not affect the methodological objective of the study, which is to evaluate how backbone extraction behaves in archaeological similarity networks constructed from survey-based ceramic data.

The dataset is modelled as a bipartite structure linking sites and ceramic categories, subsequently projected into a site–site similarity network using binary presence/absence data. Chronologically informative ceramic classes such as Red-Black Burnished Ware, Plain Simple Ware, Slipped Ware, and more selectively distributed classes provide the relational basis for measuring similarity.

The Amuq Valley constitutes a region characterised by long-term settlement continuity and broad technological integration during the Bronze Age (Casana and Wilkinson 2005; Wilkinson, 2003). As detailed in Section 2.3, its Bronze Age ceramic record is dominated by a few widely diffused classes alongside a small number of rare ones, a combination that makes the dataset particularly suitable for evaluating how similarity metrics behave under conditions of regional cultural integration and low typological diversity.

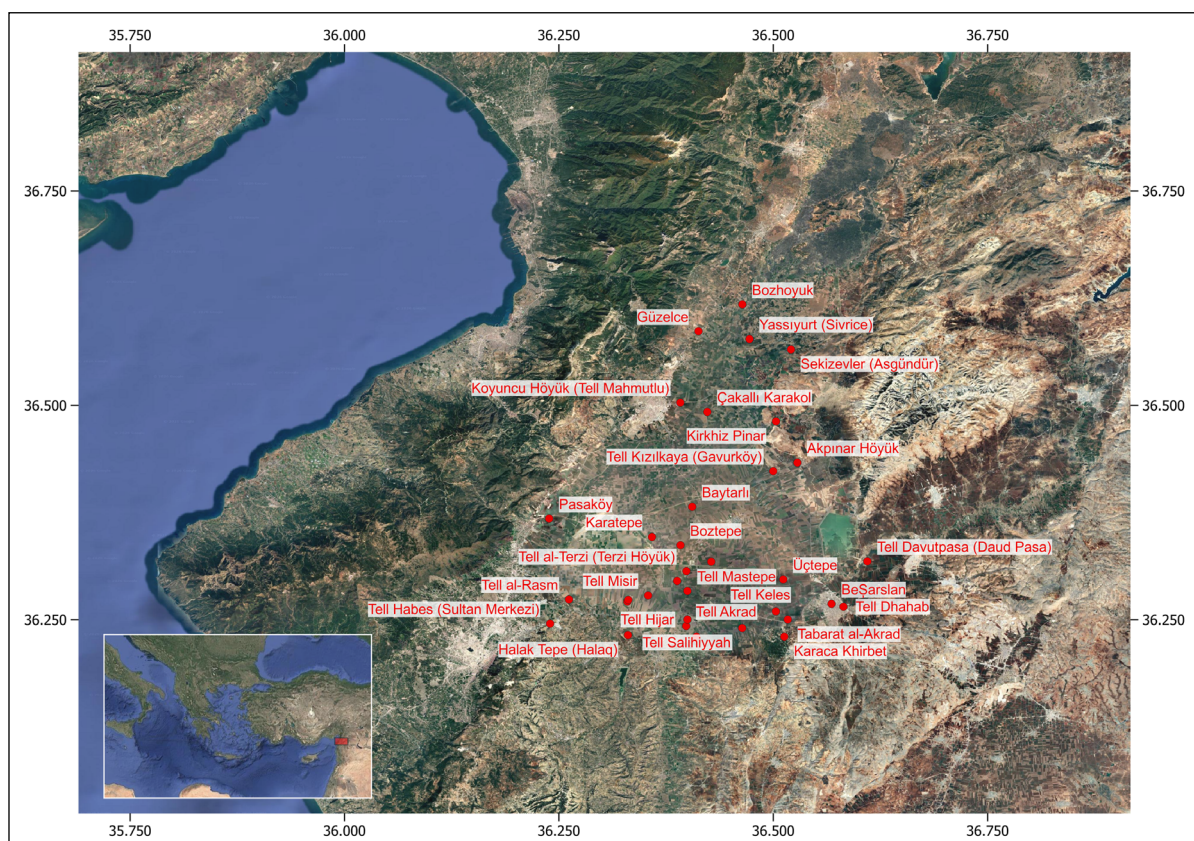


Figure 1 Location of Bronze Age sites analysed in the Amuq Valley. Regional map showing the spatial distribution of the 35 surveyed sites included in the dataset used for the similarity network analysis.

SITE	RED-BLACK BURNISHED WARE	PLAIN SIMPLE WARE	WHITE SLIPPED WARE	SLIPPED WARE
Guzelce	1	1	0	0
Yassiyurt (Sivrice)	0	0	1	0
Pasakoy	1	1	0	0
Koyuncu Hoyuk (Tell Mahmutlu)	1	1	0	0
Kirkhiz Pinar	1	0	0	0
Tell Kizilkaya	1	1	0	0
Tell Baytarli (Toprakli)	0	0	1	0
Akpınar Hoyuk	1	1	0	0
Tell Misir	1	0	0	0
Tell al-Rasm	0	1	0	0
Tell Uzunarab (Bozhoyuk)	1	1	0	0
Karatepe	1	1	0	0
Boztepe	1	1	0	0
Tell Hasanusagi	1	1	0	1
Tell al-Terzi (Terzi Hoyuk)	1	0	0	0
Kucuk Avara (Turhan Bey Ciftlik)	1	0	0	0
Tell Keles	1	1	0	1
Tulul Salihyyah al-Saghir	1	0	0	0
Halak Tepe (Halaq)	1	1	0	0
Tell Akrad	1	0	0	0
Tell Saluq	1	0	0	0
Besarlan (Tell Hamda)	1	1	0	0
Tell Mastepe (Mastepe)	1	0	0	0
Tell Davutpaia (Daud Pasa)	1	0	0	0
Putoglu (Uctepe)	1	1	0	0
Karaca Khirbet 'Ali	1	0	0	0
Tell Ermeneia	1	1	0	1
Tell Dhahab (Altin Tepe)	1	0	0	0
Tell Hajar	1	0	0	0
Tabarat al-Akrad (Tell al-Hayey)	1	0	0	0
Temel Kizilkaya	1	0	0	0
Sekizevler (Asgundur)	1	1	0	0
Kucuk Avara (South Çiftlik)	0	1	0	0
Tell Habes (Sultan Merkezi)	1	0	0	0
Cakalli Karakol	1	0	0	0

Rather than serving as a comprehensive regional reconstruction, the dataset is treated here as a controlled empirical framework for assessing the inferential behaviour of backbone extraction in archaeological similarity networks.

2.2 SURVEY BIAS AND TIME-AVERAGING

The AVR data derive primarily from systematic field survey. As with most survey-based assemblages, ceramic materials are cumulative and subject to post-depositional processes, including erosion, redeposition,

and differential visibility (Bintliff et al. 2007). Individual sites therefore represent aggregated material signatures rather than temporally discrete occupation phases.

Chronological attribution for many contexts is necessarily broad, and Bronze Age assemblages frequently compress Early, Middle, and Late components within single analytical units. This phenomenon of time-averaging (Bailey 2007) has direct implications for similarity analysis. Shared ceramic categories may reflect overlapping participation in long-lived technological traditions rather than contemporaneous interaction.

In such contexts, similarity networks do not map discrete social ties but represent structured fields of material commonality shaped by cumulative deposition and regional transmission processes. Any analytical approach must therefore account for the possibility that observed connectivity arises from aggregation and diffusion rather than selective relational dynamics.

2.3 MATERIAL HOMOGENEITY AND WORST-CASE CONDITIONS

The four Bronze Age ceramic classes in the dataset exhibit markedly unbalanced distributions across the 35 sites. Two classes are broadly attested: Red-Black Burnished Ware occurs at 31 of 35 sites (89%) and Plain Simple Ware at 17 (49%), reflecting shared technological repertoires and long-term cultural integration. The other two are rare: Slipped Ware is present at only three sites (Tell Hasanusagi, Tell Keles, and Tell Ermeneia) and White Slipped Ware at only two (Yassiyurt and Tell Baytarlı). This imbalance has a direct analytical consequence. Because two near-ubiquitous classes are shared by most sites, binary co-occurrence mechanically inflates pairwise similarity across the network, compressing its weight distribution; the two rare classes, by contrast, contribute almost no discriminating signal and, in the case of White Slipped Ware, link only an isolated pair of sites. The dataset is thus dominated by a small number of widely diffused categories, with localized variation confined to classes too rare to support relational inference.

This four-class resolution is not an analytical simplification but a property of the AVR P survey record itself. As Casana (2009) notes, Late Bronze Age assemblages in the Amuq are overwhelmingly dominated by undifferentiated plain wares — the standard sand-tempered buff fabrics that account for some 80–90% of the material — and can be securely distinguished from earlier second-millennium phases only through the rare imported or imitation fine wares (Mycenaean, Cypriot, Nuzi) that are almost entirely absent from surface collections. The coarse-grained, low-diversity structure of the data is therefore an empirical condition of survey-based ceramic recording in the region, not an artefact of our analytical choices. Far from being incidental, this is the very condition under which we test whether relational selectivity can be recovered: the near-ubiquity

of a few classes mechanically inflates pairwise similarity, and it is this inflation — rather than any conservatism of the binary coding — that the backbone procedure is designed to interrogate, with the consequences for relational inference taken up in Section 5.6.

When modelled using binary presence/absence data and projected from a bipartite structure, such configurations tend to produce highly dense similarity networks with compressed weight variability. Under these conditions, similarity is mechanically inflated by the diffusion of common categories, and relational variance is reduced.

From a network-analytical perspective, this constitutes a worst-case scenario for detecting selective relational structure. Dense connectivity is expected regardless of whether specific sites participated in preferential interaction. The Amuq dataset therefore functions as a methodological stress test: if backbone extraction identifies statistically selective relations under these constraints, they may be interpreted as departures from generalized diffusion. Conversely, minimal or null backbones delineate the empirical limits of inference from similarity data.

In this framework, the case study does not aim to reconstruct interaction networks, but to evaluate whether relational selectivity is recoverable at all once survey bias, time-averaging, and material homogeneity are taken into account.

3. METHODS

3.1 SIMILARITY NETWORKS AND MATERIAL CULTURE IN ARCHAEOLOGY

Similarity networks linking archaeological contexts on the basis of shared artefact categories are commonly derived from binary data and from one-mode projections of bipartite structures (Brughmans et al., 2016; Knappett, 2011; Peeples 2019). These modelling choices tend to produce highly dense graphs with compressed weight variability, in which numerous edges reflect mechanically generated similarities rather than structurally distinctive relations.

The term “similarity network” covers a broad family of models, differing in the entities treated as nodes, the features compared, and the coefficient used (Östborn and Gerding 2014); within archaeology alone these range from artefact co-occurrence and stylistic-similarity networks to assemblage-based consumption networks (e.g. Mills 2016; Moreno-Navarro 2024; Blair 2024), and comparable constructions recur across the JCAA literature. The definition adopted here — sites linked by shared artefact categories, derived from a one-mode projection of a bipartite site-by-category matrix (Peeples and Roberts 2013; Prignano et al. 2017) — is therefore one specific member of this family, and our claims about

density and weight compression are intended to apply to this construction rather than to similarity networks in general.

Archaeological similarity networks in this study are derived from one-mode projections of bipartite site-by-category matrices. We represent material culture using binary presence/absence data. Recording only whether a class is present rather than how abundant it is sidesteps interpretive biases related to production intensity, recovery rates, or surface visibility typical of survey contexts. We stress, however, that this is not a conservative choice with respect to similarity — a point on which we depart from the common assumption that presence/absence coding is inherently cautious. As set out in Section 2.3, with only four ceramic classes, two of them near-ubiquitous, binary co-occurrence strongly inflates pairwise site similarity, since most site pairs share at least one widely diffused class. Binarisation therefore tempers visibility-driven biases while simultaneously amplifying connectivity, and it is the latter effect that motivates the formal filtering introduced below. The coarse-grained categorical structure of the data therefore tends to compress the weight distribution and to overstate the apparent connectedness of the network. Since one-mode projections inherently inflate connectivity and compress weight variability. This is a risk in archaeological network analysis where dense similarity networks can be mistaken for genuine selectivity rather than background diffusion. Consequently, we introduce backbone extraction as a necessary formal procedure to distinguish statistically significant similarity from background homogeneity.

3.2 BACKBONE EXTRACTION FRAMEWORK

To address the density of projected networks, we utilize backbone extraction—a statistical filtering procedure that isolates edges carrying significant relational information while preserving the overall organization of the system (Liebig & Rao, 2016; Yassin et al., 2025). Within network science, the backbone is defined as the subset of connections that cannot be explained solely by graph density or weight distribution.

Methodological literature distinguishes between structural and statistical approaches. Structural methods (e.g., Maximum Spanning Tree, High Saliency Skeleton) select edges based on topological criteria but do not explicitly distinguish statistically informative links from those generated by random or aggregative processes (Serrano et al., 2009; Neal, 2022). In contrast, statistical methods—such as the Disparity, Noise-Corrected, or Global Statistical Significance (GloSS) filters—compare observed data against a null model. This counterfactual hypothesis assesses whether an observed pattern results from a random process compatible with selected baseline properties of the system filter (Neal, 2022; Yassin et al., 2025).

The choice of null model is critical for archaeological datasets, which are frequently characterized by skewed distributions and incomplete sampling. Imposing parametric theoretical models based on abstract distributions risks introducing unjustified assumptions about the processes behind material culture. Consequently, this study prioritizes empirical null models constructed directly from the observed distribution of edge weights. The concept of the null model occupies a central position in statistical backbone extraction methods. In general terms, a null model represents a counterfactual hypothesis that allows assessment of whether an observed pattern can be explained as the result of a random process compatible with selected baseline properties of the system (Tumminello et al., 2011). However, the choice of a null model is not neutral and has profound interpretative implications. By defining random behavior through sampling empirical variability rather than an a priori law, this approach reduces theoretical bias and aligns the analysis with the taphonomic properties of the data.

3.3 TECHNICAL IMPLEMENTATION

We implemented the Global Statistical Significance (GloSS) filter, a statistical backbone extraction method based on a global empirical null model (Coscia & Neffke, 2017).

In plain terms, this null model asks what the network would look like if each connection's strength were unrelated to which sites it joins: the observed edge weights are repeatedly reshuffled at random across the fixed set of edges, and an edge is judged significant only if its real weight is higher than would be expected under that random reassignment. Calculations were performed using a Monte Carlo approximation (Coscia & Neffke, 2017). During the simulation, network topology (nodes and edges) remained fixed, while edge weights were reshuffled according to the global empirical null model. The observed weight distribution served as the reference distribution.

Weights were randomly reassigned with replacement across the fixed set of $m = 994$ edges, effectively removing structural associations between topology and weight magnitude. The simulation involved $n = 1000$ resamplings with a fixed random seed (the `set.seed()` command fixes the random-number generator so that the simulation yields identical results on every run). For each edge, an empirical p-value was computed, expressing the probability of observing a weight equal to or greater than the empirical one, under the null hypothesis. Final edges were retained based on a statistical threshold of $\alpha = 0.05$ applied to FDR-adjusted empirical p-values.

As a complementary, locally based procedure, we also implemented the Disparity Filter (Serrano et al. 2009), which evaluates edge significance against a local rather than a global null model. For each node of degree k , the

weights of its incident edges are normalised to sum to unity ($p_{ij} = w_{ij}/s_i$, where s_i is the node strength), and each normalised weight is compared against the null expectation that a node's strength is distributed uniformly at random across its connections. Under this model the probability of observing a normalised weight at least as large as p_{ij} is given by $\alpha_{ij} = 1 - (1 - p_{ij})^{(k-1)}$; an edge is retained when α_{ij} falls below the significance level α . Because each edge is assessed independently from the perspective of the two nodes it connects, it is retained when it satisfies the significance criterion for at least one of them (Serrano et al. 2009). A special case concerns nodes of degree one: following Serrano et al. (2009), an edge incident on a degree-one node is retained only if it beats the threshold from the perspective of its other endpoint. When both endpoints have degree one, the edge has no local weight distribution on either side against which significance can be assessed, and it is therefore excluded from the backbone rather than retained by default. The filter was implemented directly in R following Serrano et al. (2009) and applied at $\alpha = 0.05$ to both the Jaccard and the cosine projections of the bipartite matrix. To confirm that the outcome was not an artefact of this particular cut-off, the analysis was repeated at $\alpha = 0.01$ and $\alpha = 0.10$; the result was stable across all three thresholds.

The entire analytical pipeline was implemented in R. To ensure full computational reproducibility, the code is provided as an RMarkdown document in the supplementary materials. The raw datasets and scripts are available via Zenodo ([10.5281/zenodo.18888360](https://doi.org/10.5281/zenodo.18888360)).

3.4 METHODOLOGICAL CONSTRAINTS AND EXCLUSIONS

The application of backbone extraction in this study is treated as a hypothesis-testing procedure rather than a purely technical operation. Its primary purpose is to evaluate whether selective relational structures can be distinguished from background material homogeneity.

Several common methodological alternatives were deliberately excluded to maintain analytical restraint. We avoided externally defined, arbitrary thresholds, which retained edges by fiat and obscured whether a tie was statistically distinguishable from background connectivity. We also set aside frequency-based weighting, since abundance schemes risked conflating surface visibility and recovery intensity with behavioural significance in survey-derived assemblages, and we did not attempt fine chronological segmentation, because the available data could not reliably support the required temporal resolution.

Community detection was likewise not applied to the full similarity network, and this exclusion warrants explanation, since clustering is a natural first instinct on a dense graph. Community detection can of course be informative on archaeological networks, but the

published cases operate on networks that are already sparse or spatially constrained rather than on a complete similarity graph: Visser (2021) detects communities on dendrochronological networks built with explicit similarity thresholds, and Peeples and Bischoff (2023) partition material-cultural networks that are strongly structured by geographic distance. Our point is not that community detection is invalid, but that on a near-complete, unfiltered similarity network dominated by ubiquitous ceramic classes it describes graded background variation rather than statistically selective ties.

Community detection and backbone extraction thus answer different questions: the former describes how a dense graph subdivides, whereas the latter tests whether any tie is statistically exceptional in the first place. Because our aim is the second question, we treat backbone extraction — not clustering — as the appropriate diagnostic here; we return to the contrast between the two procedures, applied to the same network, in Section 4.2.

4. RESULTS FROM THE CERAMIC SIMILARITY NETWORKS IN THE AMUQ VALLEY DURING THE BRONZE AGE

4.1 NETWORK STRUCTURE AND SIMILARITY DISTRIBUTIONS

On this basis, the data were modelled as a bipartite structure connecting sites and ceramic classes. The similarity network between sites was then obtained through one-mode projection using the Jaccard index, which is particularly suitable for measuring similarity between binary sets. This index normalises the sharing of ceramic types relative to the total number of categories present at each site, reducing the distorting effect of richer—or better-documented—contexts. The Jaccard index has an established record in archaeological similarity studies built on survey-derived categorical data, including rock-art assemblages where it is used to construct site-to-site similarity networks (Riris and Oliver 2019) and to normalise theme co-occurrence networks (Intxaurbe 2026). In several of these applications a threshold is applied to the resulting similarity values to reduce over-plotting and isolate stronger relations — for instance retaining only ties with Jaccard ≥ 0.1 (Riris and Oliver 2019) — which illustrates the common archaeological practice of pruning dense similarity matrices prior to interpretation, and motivates the more principled, statistically grounded filtering adopted here.

The resulting network is highly dense, with a large number of edges and a distribution of weights compressed toward medium-to-high values (Figure 2a). This configuration, typical of similarity networks derived from bipartite projections, makes direct interpretation problematic, as many relationships may emerge as a

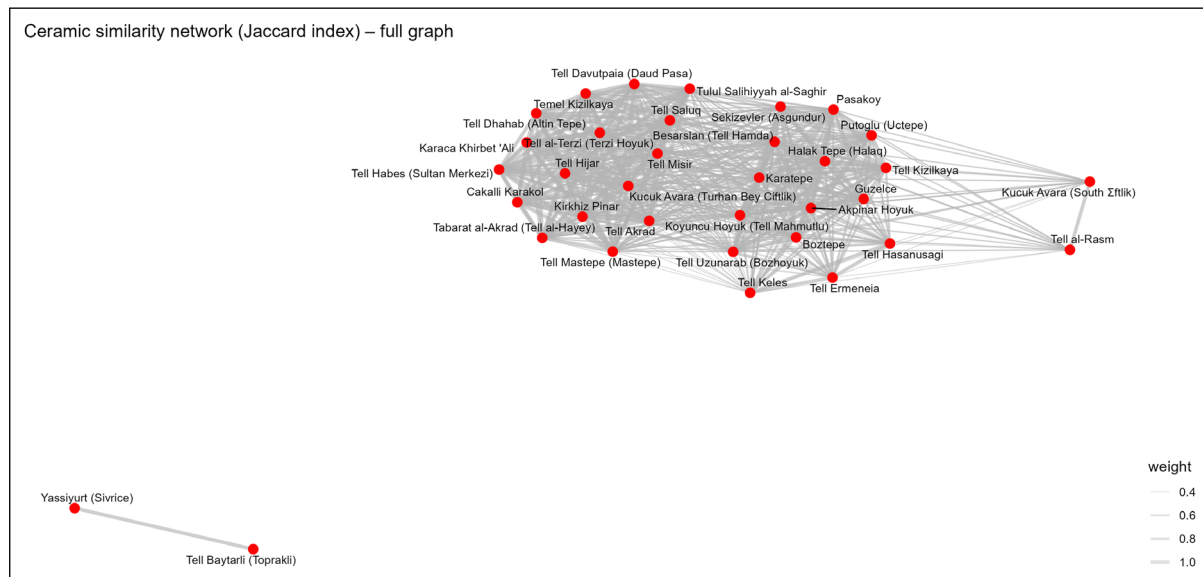


Figure 2a Full similarity network derived from ceramic co-occurrence. Network representation of the site-site similarity structure based on the Jaccard index calculated from the bipartite site-by-ceramic category matrix. Nodes represent archaeological sites and edges represent pairwise similarity relations.

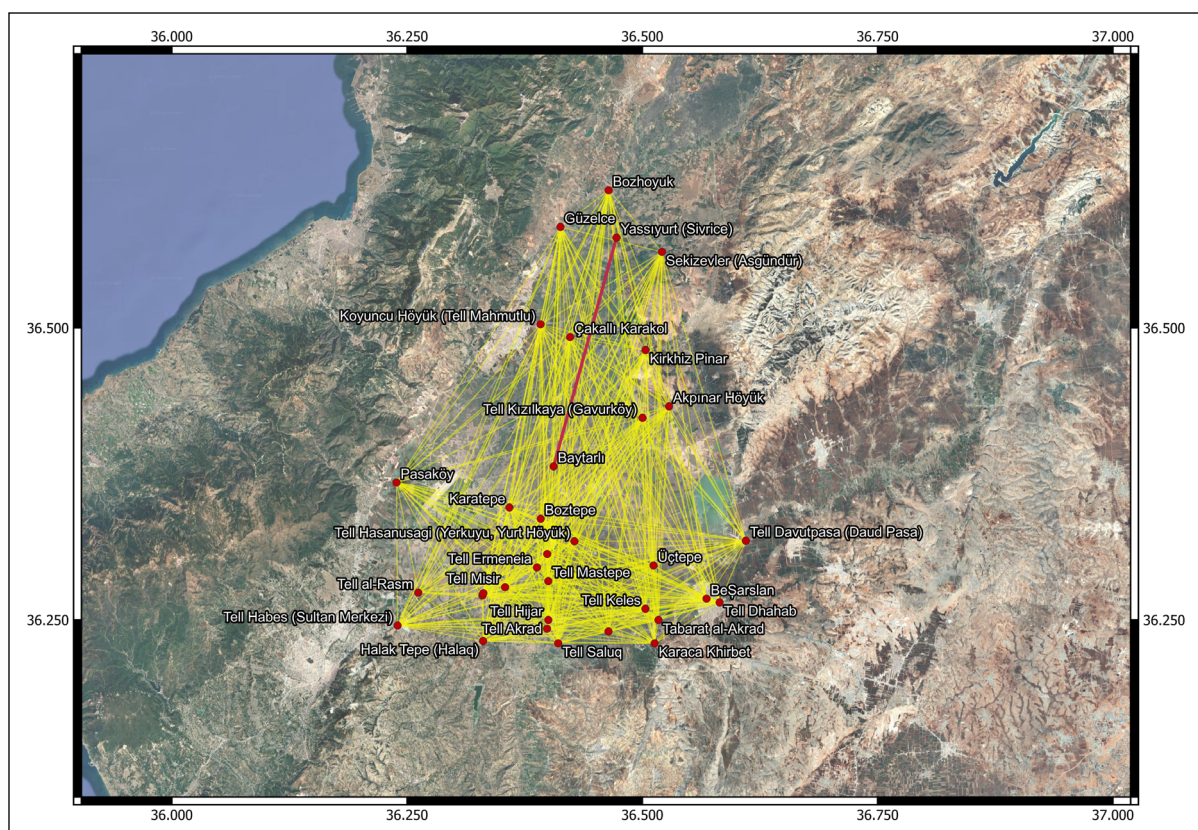


Figure 2b Geographic projection of the complete site-site similarity network over satellite imagery. Nodes are placed at their real-world coordinates in the Amuq Valley; yellow edges represent pairwise Jaccard similarities. Near-complete connectivity spans the full spatial extent of the surveyed sites without forming spatial clusters, indicating regional-scale diffusion rather than localized interaction. The red edge marks the only exclusive tie — Yassiyurt (Sivrice)–Tell Baytarli (Toprakli), sharing White Slipped Ware (cf. Section 5.3) — which carries the same weight as all others despite its rarity.

mechanical effect of the regional diffusion of widely shared ceramic classes rather than as indicators of historically meaningful relationships between sites (Figure 2b).

With 35 sites and nearly complete connectivity, the network approaches the upper bounds of density

expected for undirected similarity graphs derived from bipartite projections. Under these conditions, differences in edge weights are constrained within a narrow interval, limiting the emergence of extreme values that could be interpreted as relationally distinctive. High

density therefore does not simply complicate visual interpretation; it directly affects the statistical behaviour of filtering techniques by reducing variance in the weight distribution.

To evaluate the robustness of these structural properties with respect to the choice of similarity measure, we compared the weight distributions obtained using the binary Jaccard index and cosine similarity derived from the same bipartite representation (Figure S2). Although the two measures differ mathematically, both produce strongly right-skewed and compressed distributions. This convergence indicates that the observed structural configuration is primarily driven by the underlying data—characterised by the widespread diffusion of a limited number of ceramic classes—rather than by the specific similarity metric employed.

These structural properties directly constrain the statistical behaviour of backbone extraction.

4.2 BACKBONE EXTRACTION OUTCOMES

In a first analytical step, the Global Statistical Significance (GloSS) filter was applied, based on a global empirical null model comparing the observed weight of each edge with a distribution of simulated weights obtained by sampling the overall weight distribution of the network. Application of the GloSS filter did not yield statistically significant edges for conventional thresholds ($\alpha \leq 0.10$) (Figure 3). This outcome reflects the conservative behaviour of global null models in dense similarity networks with compressed weight distributions, rather than a methodological failure.

In light of this result, a local filtering approach based on the Disparity Filter was adopted. This method evaluates edge significance at the node level, comparing the normalised weight of each link with the distribution of weights incident on each node (Serrano et al. 2009). The Disparity Filter was applied to both the Jaccard

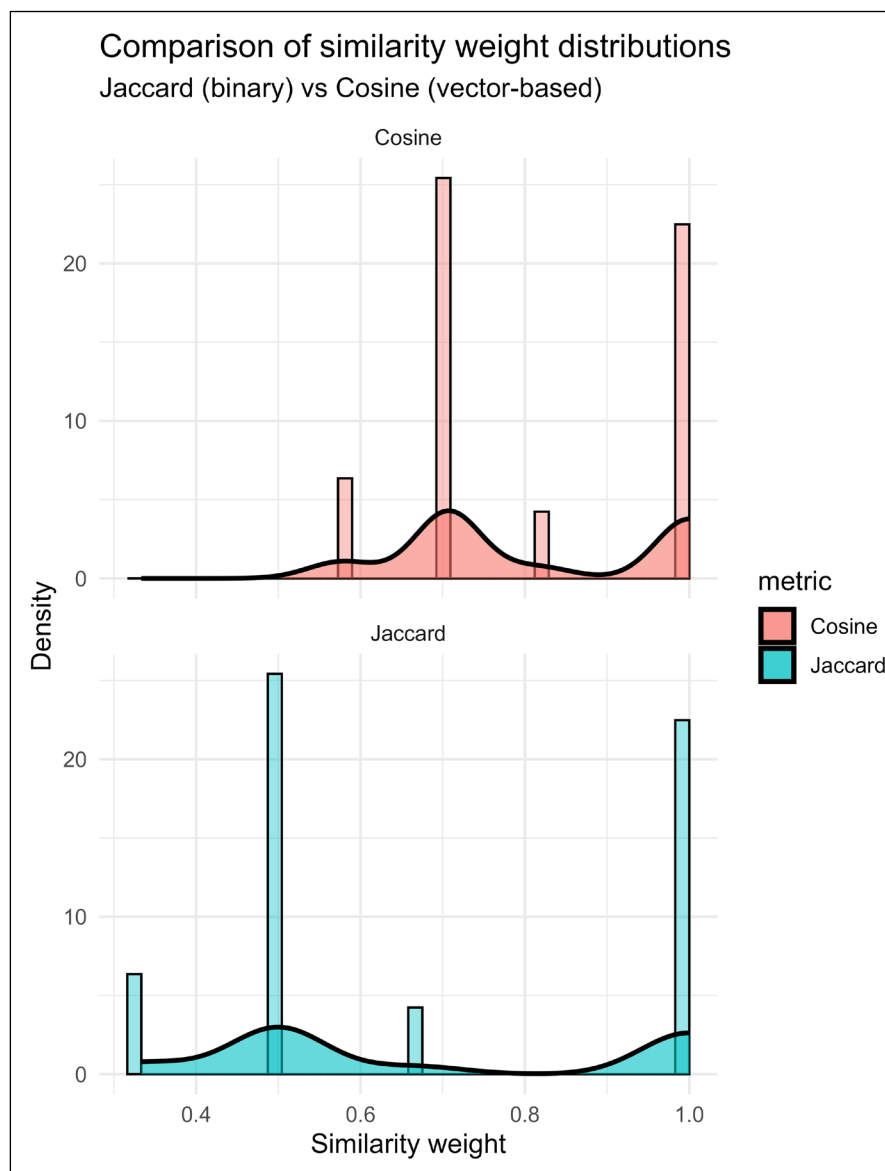
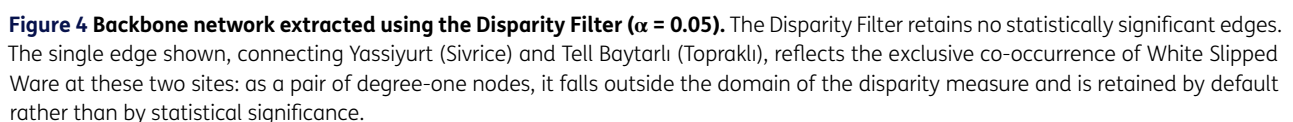
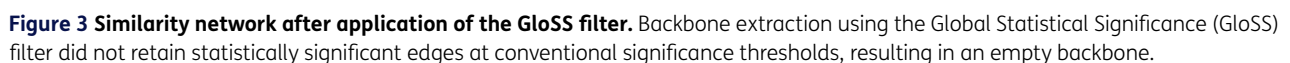


Figure S2 Comparison of similarity weight distributions for Jaccard and cosine metrics. Distribution of similarity weights derived from the Jaccard index (binary co-occurrence) and cosine similarity (vector-based measure) calculated from the same bipartite dataset.

be assessed as statistically selective, and its persistence reflects a structural property of the data rather than a recovered relational signal.

From a dataset perspective, the White Slipped Ware class is attested exclusively at two sites: Yassiyurt (Sivrice) and Tell Baytarlı (Topraklı). These two contexts therefore connect exclusively to one another, sharing a category that is not explained by the general diffusion of the dominant ceramic classes. Within the projected network, however, this exclusivity translates into a pair of degree-one nodes, for which the disparity measure is undefined: a node with a single incident edge has no local weight



distribution against which significance can be assessed (Serrano et al. 2009).

The two sites are nonetheless of archaeological interest in their own right. Geographically, they are located within the same macro-sector of the Amuq plain, at a distance compatible with direct interaction or participation in a shared catchment or procurement basin. Chronologically, both present coherent attestations of Middle–Late Bronze Age ceramics as documented by the AVR P survey, and the co-presence of Red-Black Burnished Ware and Plain Simple Ware points to technological and functional convergence. These observations remain descriptively meaningful, but they cannot be promoted to a statistically validated relation: the backbone procedure neither confirms nor rejects a tie between contexts that the filter is structurally unable to evaluate.

Importantly, the surviving edge should not be read as a statistically selected relation. It is retained by default rather than by significance, and is therefore excluded from the backbone. The negative outcome of the GloSS filter and the absence of any statistically validated tie under the Disparity Filter together demonstrate that backbone extraction here, functions as a diagnostic tool: it tests whether selective relational structure can be distinguished from background material diffusion, and in this dataset it cannot.

To make the contrast in Section 3.4 concrete, we applied a standard community-detection algorithm (Louvain modularity optimisation) to the full Jaccard similarity network. The algorithm returns a partition of the 35 sites into three communities (modularity $Q = 0.17$), and this partition is stable across runs. Yet it coexists with an empty statistical backbone: the same network from which no single edge survives significance filtering is nonetheless divisible into communities. The two results are not in contradiction — they measure different things. The modularity value is low, and the partition largely reflects graded variation in the presence of the two dominant, near-ubiquitous wares across the plain rather than statistically selective relations between specific sites; tellingly, the only sharply separated group is the two-site community defined by the exclusive co-occurrence of White Slipped Ware, the same degree-one configuration discussed above. The presence of a modular partition therefore does not, by itself, demonstrate relational selectivity: community detection will always return communities from a connected graph, whereas only the backbone test asks whether any tie is statistically exceptional — and here it returns a negative result.

The absence of an articulated backbone should therefore be interpreted as a fully fledged empirical result. It indicates that most observed ceramic similarities are consistent with a broadly shared technological and cultural horizon — rather than positively demonstrating discrete interaction networks — a reading whose interpretive limits are discussed in Section 5.5. Under

conditions of minimal information and maximum redundancy—typical of survey-based binary datasets—backbone extraction isolates only punctual relations driven by rarity rather than intensity.

Future extensions could explore weighted similarity measures or probabilistic projections; however, the present binary framework intentionally tests backbone behaviour under conservative informational conditions.

4.3 SECOND CASE STUDY: THE CARPETANIA CONSUMPTION NETWORK

To test whether the empty backbone reflects the structure of survey-derived ceramic data rather than the way the Amuq network is constructed, we applied the identical analytical pipeline to a second, independent dataset: the Roman-period rural consumption network of northern Carpetania (Moreno-Navarro 2024). This dataset was chosen because it is structurally the opposite of the Amuq in the dimension most likely to drive a null result. Where the Amuq matrix records four ceramic wares, the Carpetania matrix records 123 chronotypes across nine sites — an order of magnitude greater typological resolution, compiled by a different author, for a different region and period, and for an unrelated research question. If the empty backbone were an artefact of low typological diversity or of our particular construction choices, this far richer dataset should behave differently. The counts were reduced to binary presence/absence to match the coding used for the Amuq, and the same GloSS and Disparity filters were applied to both the Jaccard and the cosine projections, at $\alpha = 0.01, 0.05$ and 0.10 .

The outcome is the same. The full Jaccard similarity network is complete (density = 1.0), and its edge weights fall within a narrow band (0.10–0.36; mean 0.24, sd 0.06), closely mirroring the compressed distribution of the Amuq network. No edge survives statistical filtering under any of the three filters at any threshold: the backbone is empty in every case.

The reason this high-dimensional dataset behaves like the four-ware Amuq case lies in the distribution of its categories. Of the 123 chronotypes, 83 occur at a single site and only three are present at all nine. Types confined to one site cannot, by construction, generate any pairwise similarity, so the similarity signal is carried by a small number of shared classes, exactly as in the Amuq. The 123 categories are therefore high-dimensional in nominal terms but low in effective relational selectivity: the assemblage combines widespread redundancy with non-shared rarity, and neither condition produces statistically distinctive ties. We note that the Carpetania network is small (nine sites, yielding a complete graph in which every node has degree eight), so the disparity filter operates close to its degenerate limit; the substantive point is therefore not the empty backbone per se, but that it arises through the same structural mechanism — weight compression driven by non-shared rarity — that governs the Amuq result.

5. DISCUSSION

The application of backbone extraction to archaeological similarity networks raises methodological and interpretative issues that extend well beyond the specific case study presented here. The Amuq Valley ceramic dataset provides an instructive testing ground, as it combines several features common to archaeological survey data: binary recording, bipartite data structures, cumulative surface assemblages, and a high degree of informational redundancy resulting from the widespread diffusion of material categories. Taken together, these features produce dense relational fields in which apparent connectivity is high, but the historical selectivity of specific ties remains difficult to assess.

5.1 BACKBONE EXTRACTION AS A DIAGNOSTIC RATHER THAN REDUCTIVE TOOL

In many network applications, backbone extraction is implicitly understood as a reductive procedure aimed at simplifying complex topologies for visualisation or interpretation (for archaeological and adjacent uses of edge-reduction and thresholding, see [Riris and Oliver 2019](#); [Neal 2014](#)). From this perspective, the analytical value of a backbone is often measured by its ability to reveal articulated structures, clusters, or hierarchies. However, this expectation risks conflating methodological success with the generation of structurally rich outputs.

The present study shifts this focus. Backbone extraction is employed not to enforce structure on dense networks, but as a diagnostic procedure to evaluate whether such structure is statistically supported. In this framework, the backbone functions as a formal hypothesis-testing mechanism, determining whether observed similarities between archaeological contexts exceed the expectations of generalised material diffusion and cumulative depositional processes.

This distinction is critical in archaeology, where similarity networks are frequently products of long-term traditions, shared technological repertoires, and overlapping chronological horizons, rather than discrete episodes of interaction. Under these conditions, dense networks should not be viewed as *prima facie* evidence for complex relational systems. Instead, they demand analytical tools capable of disentangling background similarity from statistically selective relationships. When grounded in explicit null models, backbone extraction offers a robust framework for this task. From this perspective, analytical utility is defined not by the resulting structural complexity, but by the method's capacity to formally constrain interpretation. In methodological terms, this study advocates for the use of backbone extraction as a hypothesis-testing device rather than a mere visualization tool. Aligning the null model with the taphonomic and empirical properties of

the data allows for a conservative and robust analytical framework, ensuring that any retained network structure represents a genuinely exceptional configuration rather than a statistical artefact.

5.2 GLOBAL VERSUS LOCAL NULL MODELS AND THE MEANING OF NEGATIVE RESULTS

From a methodological perspective, the divergent outcomes produced by the Global Statistical Significance (GloSS) filter and the Disparity Filter highlight the importance of aligning backbone extraction techniques with the structural properties of the data and the broader analytical aims. The inability of the GloSS filter to identify statistically significant ties should be understood not as a methodological failure but as a direct consequence of the model's conservative assumptions. This outcome reflects the inherent difficulty global null models face when applied to similarity networks with compressed weight distributions and high redundancy.

In the Amuq Valley network, most similarity values fall well within the range predicted by empirical resampling. Consequently, global comparison treats the majority of relationships as consistent with average system behaviour, suppressing all candidate edges. This outcome is analytically significant, as it indicates that inter-site similarity is primarily structured by regional-scale processes rather than by selective or exceptional relationships.

The Disparity Filter, which evaluates edge significance at a local level relative to the individual node profiles, might in principle be expected to recover site-specific deviations that a global model overlooks. In the Amuq network, however, it too returns an empty backbone. This convergence is consistent with the conditions under which the disparity filter was designed to operate: [Serrano et al. \(2009\)](#) note that the method exploits strong disorder in the weight distribution, both globally and locally, and that where heterogeneity is low and weights are not locally correlated the procedure loses its discriminating power and behaves much like a global threshold. The Amuq similarity network, built from a small number of broadly diffused binary categories, is precisely such a low-disorder case. The absence of a recoverable backbone is therefore not an artefact of an inappropriate method, but the expected behaviour of local filtering when the underlying data carry little relational heterogeneity.

5.3 RARITY WITHOUT RECOVERABLE SELECTIVITY

The case of the connection between Yassiyurt (Sivrice) and Tell Baytarlı (Topraklı) is particularly informative when contextualized within the dataset's structure, precisely because it does not yield a statistically significant tie. The elusive presence of White Slipped Ware at only two sites

might intuitively appear to signal a relationally specific link. Yet this exclusivity produces a pair of degree-one nodes, a configuration that lies outside the domain of statistical backbone extraction: there is no local weight distribution against which the disparity measure can assess significance, and no global contrast that the GloSS filter can detect.

In this case, the surviving edge does not represent a quantitatively stronger relationship in absolute terms. It reflects a similarity that cannot be attributed to the mechanical effect of widespread ceramic diffusion, since it rests on a category shared by these two sites alone. Yet this very exclusivity is what places it beyond the reach of statistical filtering: the co-occurrence of a rare category between two otherwise isolated nodes deviates from the general homogeneity of the network, but it cannot be validated as a selective tie, because the filters have no local or global baseline against which to evaluate it.

This observation carries broader implications for archaeological network research. In datasets dominated by a few ubiquitous categories, the signals most likely to appear relationally specific are the rarest ones — yet these are precisely the signals that tend to generate degree-one configurations, which statistical filters cannot validate as selective. Rarity alone, in other words, does not translate into statistically recoverable selectivity. Rather than promoting such uncommon material signals to markers of interaction, backbone extraction correctly declines to validate them, foregrounding instead the limits of what survey-derived binary data can support. The White Slipped Ware case thus exemplifies, rather than contradicts, the central result: under conditions of minimal typological information and high redundancy, selective relational structure is not recoverable, whether through frequency or through rarity.

5.4 THE “EMPTY BACKBONE” AS AN ARCHAEOLOGICAL RESULT

The extreme sparsity of the extracted backbone should not be interpreted as a methodological failure, but as a significant analytical finding. In computational archaeology, there is often a bias toward producing visually complex and “structured” networks. However, forcing structure onto a dataset that lacks relational selectivity leads to over-interpretation and “false positives” in historical reconstruction.

The collapse of the network into a dense field with few significant edges reinforces the need for interpretive restraint. Our findings indicate that, for the period and region analyzed, material culture was not organized into the kind of exclusive, highly selective clusters that typical network algorithms (such as community detection) are designed to find. This high degree of “entrenchment” of ceramic types suggests a system where material signals are too blurred to support strong claims about discrete social boundaries or closed exchange circuits. In this

sense, the empty backbone does not signal analytical weakness, but rather the structural homogeneity of the underlying material system.

5.5 SIMILARITY, INTERACTION, AND THE LIMITS OF INFERENCE

A persistent interpretative risk in archaeological network analysis lies in the tendency to equate material similarity with social interaction. Networks of shared traits are often treated, implicitly or explicitly, as proxies for exchange or communication. The present study deliberately avoids such an equivalence.

Ceramic similarity may emerge from a wide range of processes, including shared technological traditions, long-term cultural convergence, and overlapping chronological horizons. In survey-derived assemblages, these processes are often temporally compressed, further blurring distinctions between synchronic and diachronic accumulation. As a result, similarity networks should be conceptualized as representations of material commonality rather than as direct maps of socioeconomic interaction.

Within this framework, the absence of an articulated backbone structure is highly informative, but its historical reading must be stated with care. The result is grounded in a specific and quantifiable property of the assemblage: the near-ubiquity of Red-Black Burnished Ware (89% of sites) and Plain Simple Ware (49%), against two classes too rare to discriminate. This distribution is consistent with shared technological repertoires and sustained low-level circulation across the Amuq during the Bronze Age, but the survey record does not allow this background commonality to be resolved into selective, site-to-site relationships. A broadly shared regional horizon is therefore one interpretation compatible with the data, rather than a conclusion the data positively demonstrate; equally compatible are cumulative time-averaging and the coarse typological resolution of the survey record (Section 2.3). Distinguishing among these would require finer-grained evidence — additional ceramic traits, quantified assemblages, or stratified contexts — that survey data do not provide. By foregrounding these limits rather than projecting structure onto them, backbone extraction constrains archaeological interpretation and explicitly defines the inferential boundaries of similarity-based data. One way to move beyond material commonality toward interaction would be to bring inter-site distance into the analysis directly. Because the two sites sharing White Slipped Ware lie within the same macro-sector of the plain, at a distance compatible with direct contact, spatial proximity is a plausible additional axis of explanation. A spatially informed model — weighting or constraining ties by geographic distance, or testing the observed similarities against a distance-decay or gravity expectation — would ask whether shared material tracks geography, a different

and complementary question to the one backbone extraction addresses. We do not pursue it here, since it would shift the paper from the recoverability question toward a spatial-interaction model, but we return to it as a direction for future work (Section 5.8).

5.6 ROBUSTNESS TO SIMILARITY MEASURES AND DATA-DRIVEN OUTCOMES

The comparison between Jaccard-based and cosine-based similarity networks further reinforces these findings. Despite differences in mathematical formulation, both measures produce strongly right-skewed and compressed weight distributions, resulting in extremely sparse backbones when subjected to statistical filtering. This convergence suggests that the outcomes are primarily driven by the data structure rather than by the choice of similarity metric. The limited number of ceramic classes is not a mere analytical simplification; it reflects the actual typological resolution of the AVRPA dataset for the Bronze Age. The structural compression of relational variance is therefore an empirical property of the archaeological record rather than a methodological artefact.

In contexts characterized by limited typological diversity and widespread material diffusion, similarity networks are expected to manifest as dense relational fields with few statistically exceptional connections. The robustness of this outcome — namely the extreme sparsity of statistically validated relations across similarity specifications — underscores the importance of treating minimal or null outcomes as data-driven findings rather than as methodological failures.

5.7 IMPLICATIONS FOR SURVEY-BASED ARCHAEOLOGICAL NETWORKS

More generally, this case study highlights the particular challenges posed by survey-based archaeological datasets for network research. Surface assemblages are inherently cumulative, spatially uneven, and often lack high-resolution chronology. Similarity networks derived from such data are therefore especially susceptible to informational redundancy and over-connectivity.

Backbone extraction offers a principled way to address these issues, not by imposing a predetermined structure, but by formally testing for its presence. When applied critically and reflexively, it provides a safeguard against over-interpretation, ensuring that any inferred relational patterns are grounded in statistically defensible deviations from null expectations.

These considerations translate into a set of practical recommendations for researchers working with similarity networks, particularly those derived from survey data. Before interpreting such a network as a relational structure, analysts should first inspect the distribution of edge weights: backbone methods based on local null models are designed for networks with strong weight heterogeneity, and a compressed or

strongly right-skewed distribution is an early indication that the filter may lose its discriminating power and behave much like a simple global threshold (Serrano et al. 2009). Second, the choice of filtering method and null model should be made explicit and matched to the analytical question, since different methods are known to yield substantially different backbones with different structural properties (Yassin et al. 2025); in particular, it matters whether deviation is assessed against the whole network (a global model such as GloSS) or against each node's local profile (a local model such as the Disparity Filter). Third, the behaviour of backbone extraction should be evaluated in light of the data structure itself — the number and balance of categories, the degree of typological redundancy, and the presence of nodes whose connections are too sparse to be assessed, such as the degree-one nodes generated by classes attested at only two sites. Where these conditions are adverse, as in binary survey assemblages built from few, broadly diffused categories, a null or minimal backbone should be anticipated and read as a substantive result rather than as a deficiency of the data or the method. Adopting these steps as a routine diagnostic stage, prior to any relational interpretation, would help guard against the recurrent tendency to over-read dense similarity networks as evidence of structured interaction.

5.8 IMPLICATIONS FOR COMPARATIVE ARCHAEOLOGICAL NETWORK STUDIES

Beyond the Amuq Valley case, these results suggest that similarity networks derived from survey-based, binary, and cumulative datasets represent structurally extreme scenarios for backbone extraction. The convergence of widespread material traditions, limited typological diversity, and chronological compression systematically inflates connectivity while compressing relational variance.

The second case study makes this argument concrete. The Amuq and Carpetania datasets sit at opposite extremes of typological resolution — four wares against 123 chronotypes — yet both yield an empty backbone under every filter and threshold, and both do so for the same structural reason: a similarity signal dominated by a few ubiquitous classes and diluted by categories too rare or too widely shared to discriminate. That two networks of such different composition, assembled independently and for different purposes, converge on the same outcome indicates that the null backbone is a property of survey-derived, binary, cumulative assemblages rather than of any single network's construction. It follows that a null or minimal backbone, obtained under these informational conditions, is not evidence that a given method has failed on a given dataset, but a recurrent and predictable behaviour that defines the boundary conditions of relational inference for this class of data. Conversely, the same procedures applied to assemblages

with greater typological resolution and stronger weight heterogeneity would be expected to recover an informative backbone; a systematic comparison across datasets of differing resolution, and the integration of spatial distance as an additional axis of analysis (Section 5.5), are promising directions for establishing where along this gradient relational selectivity becomes statistically recoverable.

In such contexts, backbone extraction is reframed: it functions not as a tool for revealing latent complexity, but as a formal stress test for the existence of relational selectivity. Rather than assuming that filtering will expose hidden structures, these datasets force a critical evaluation of whether such structures are statistically present once the effects of diffusion and aggregation are explicitly modelled.

Consequently, minimal or null backbones emerge as epistemically productive results that define the boundary conditions of inference. This approach provides a comparative framework for distinguishing contexts where relational selectivity is statistically recoverable from those dominated by structural homogeneity. Ultimately, backbone extraction contributes to a more restrained and empirically grounded theory of material connectivity, ensuring that interpreted patterns are substantiated by the statistical properties of the archaeological record.

6. CONCLUSIONS

This study demonstrates that backbone extraction provides a methodologically robust framework for evaluating archaeological similarity networks, particularly those characterised by high density and informational redundancy. By applying statistical filtering grounded in empirical null models, this approach allows for a formal assessment of whether observed similarities reflect selective relational patterns or are instead consistent with generalised processes of cultural diffusion.

The Amuq Valley case shows that, when similarity networks are derived from binary presence/absence data and ubiquitous material categories, backbone extraction may not necessarily produce complex topologies. Instead, it produces an empty backbone across both global and local filters, indicating that no inter-site similarity exceeds the expectations of generalised material diffusion. In this respect, the empty backbone is not an arbitrary simplification, but a formal demonstration that no relationship exceeds global or local null expectations.

Crucially, the absence of an articulated backbone across similarity measures should be interpreted not as a methodological limitation, but as a substantive empirical result. In the present case, this outcome is consistent with — rather than positive proof of — ceramic similarities across the Amuq Valley reflecting a broad regional

horizon rather than discrete spheres of interaction; as discussed in Section 5.5, this is one interpretation compatible with a homogeneous, low-resolution survey record, and the data do not allow it to be resolved into selective relationships. These findings underscore the necessity of a critical and reflexive application of network analysis in archaeology: backbone extraction serves as a powerful diagnostic instrument to test the selectivity of inferred relations, preventing the over-interpretation of dense similarity structures.

More broadly, this study contributes to an ongoing reconsideration of the inferential limits of network-based methods in archaeological research. It emphasises that methodological rigour does not necessarily lead to increasingly complex representations; rather, it may instead produce constrained or minimal outcomes that define clearer boundaries for archaeological interpretation. Such results are analytically robust, ensuring that the study of past connectivity remains firmly grounded in the statistical properties of the material record.

In this respect, backbone extraction should be conceptualized as a transferable evaluative framework rather than as a technique tied to a specific materials or region. Applied comparatively, it provides a formal mean to assess the extent to which similarity patterns are shaped by selective relational processes as opposed to generalised cultural diffusion. Crucially, the absence of selective relational structure does not imply a lack of social processes; rather, it indicates that such processes are not recoverable from similarity data alone. This shift—from descriptive mapping of connectivity to formal testing of its statistical selectivity—represents a substantive methodological contribution to the application of network science in archaeology.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE MANUSCRIPT PREPARATION PROCESS

During the preparation of this manuscript, the authors used AI-assisted language tools for translation and stylistic revision. All content was critically reviewed and edited by the authors, who take full responsibility for the accuracy and integrity of the final version of the work.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

ACKNOWLEDGEMENTS

The authors wish to thank Marina Pucci for her valuable comments on an earlier draft of this manuscript.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

Dario Nincheri: Conceptualization, methodology, data analysis, and results interpretation.

Marco Moderato: Methodology validation and critical review, literature search (state of the art), and drafting of the discussion.

All authors have contributed to drafting or revising the manuscript, approved the final version to be published, and agree to be accountable for all aspects of the work in ensuring its accuracy and integrity.

AUTHOR AFFILIATIONS

Dario Nincheri  orcid.org/0009-0008-2936-6811

Universidad de Cadiz, ES

Marco Moderato  orcid.org/0000-0002-7942-3385

Universitat Pompeu Fabra, ES

REFERENCES

- Akar, M** and **Kara, D.** 2018. 'Into the Hinterland: The Middle Bronze Age Building at Toprakhisar Höyük, Altınözü (Hatay, Turkey)'. *Adalya*, 21: 81–116.
- Bailey, G.** 2007. 'Time perspectives, palimpsests and the archaeology of time'. *Journal of Anthropological Archaeology*, 26(2): 198–223. DOI: <https://doi.org/10.1016/j.jaa.2006.08.002>
- Bintliff, JL, Howard, P** and **Snodgrass, AM.** 2007. *Testing the Hinterland: The Work of the Boeotia Survey (1989–1991) in the Southern Approaches to the City of Thespiæ*. Cambridge: McDonald Institute for Archaeological Research.
- Blair, EH.** 2024. 'Material Culture Similarity and Co-occurrence Networks'. In: Brughmans, T, Mills, BJ, Munson, J and Peeples, M (eds.) *The Oxford Handbook of Archaeological Network Research*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780198854265.013.6>
- Braidwood, RJ.** 1937. *Mounds in the Plain of Antioch: An Archaeological Survey*. Chicago: University of Chicago Press.
- Brughmans, T, Collar, A** and **Coward, F.** 2016. *The Connected Past: Challenges to Network Studies in Archaeology and History*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oso/9780198748519.001.0001>
- Brughmans, T, Mills, BJ, Munson, J** and **Peeples, M.** (eds.) 2024. *The Oxford Handbook of Archaeological Network Research*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780198854265.001.0001>
- Brughmans, T** and **Peeples, M.** 2017. 'Trends in archaeological network research: A bibliometric analysis'. *Journal of Historical Network Research*, 1: 1–24. DOI: <https://doi.org/10.25517/jhnr.v1i1.10>
- Brughmans, T** and **Peeples, MA.** 2023. *Network Science in Archaeology*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/9781009170659>
- Casana, J.** 2009. 'Alalakh and the archaeological landscape of Mukish: The political geography and population of a Late Bronze Age kingdom'. *Bulletin of the American Schools of Oriental Research*, 353: 7–37. DOI: <https://doi.org/10.1086/BASOR27805145>
- Casana, J, Gansell, AR** and **Yener, KA.** 2005. Surface ceramics, off-site survey, and floodplain development at Tell Atchana (Alalakh). In: Yener, KA (ed.) *The Amuq Valley Regional Projects 1: Surveys in the Plain of Antioch and Orontes Delta, Turkey, 1995–2002*. Chicago: Oriental Institute of the University of Chicago, pp. 199–224.
- Casana, J** and **Wilkinson, TJ.** 2005. Settlement and landscapes in the Amuq region. In: Yener, KA (ed.) *The Amuq Valley Regional Projects 1: Surveys in the Plain of Antioch and Orontes Delta, Turkey, 1995–2002*. Chicago: Oriental Institute of the University of Chicago, pp. 25–65.
- Collar, A, Coward, F, Brughmans, T** and **Mills, BJ.** 2015. 'Networks in archaeology: Phenomena, abstraction, representation'. *Journal of Archaeological Method and Theory*, 22(1): 1–32. DOI: <https://doi.org/10.1007/s10816-014-9235-6>
- Coscia, M** and **Neffke, F.** 2017. Network backboning with noisy data. In: *Proceedings of the 33rd IEEE International Conference on Data Engineering (ICDE)*. DOI: <https://doi.org/10.1109/ICDE.2017.100>
- Donnellan, L.** (ed.) 2020. *Archaeological Networks and Social Interaction*. London: Routledge. DOI: <https://doi.org/10.4324/9781351003063>
- Intxaurbe, I.** 2026. 'Mapping the symbolic structure of Palaeolithic rock art using co-occurrence network analysis'. *Journal of Archaeological Method and Theory*, 33: 59. DOI: <https://doi.org/10.1007/s10816-026-09796-y>
- Knappett, C.** 2011. *An Archaeology of Interaction: Network Perspectives on Material Culture and Society*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:osobl/9780199215454.001.0001>
- Liebig, J** and **Rao, A.** 2016. 'Fast extraction of the backbone of projected bipartite networks to aid community detection'. *Europhysics Letters*, 113(2): 28003. DOI: <https://doi.org/10.1209/0295-5075/113/28003>
- Mills, BJ.** 2016. Communities of Consumption: Cuisines as Constellated Networks of Situated Practice. In: *Knowledge in Motion: Constellations of Learning Across Time and Place*. Amerind Studies in Anthropology. Tucson, AZ: University of Arizona Press. pp. 247–270.
- Moreno-Navarro, F.** 2024. 'Similarity in consumption patterns among peasant communities in Roman Central Hispania through network science'. *Journal of Computer Applications in Archaeology*, 7(1). DOI: <https://doi.org/10.5334/jcaa.120>

- Neal, Z.** 2014. 'The backbone of bipartite projections: Inferring relationships from co-authorship, co-sponsorship, co-attendance and other co-behaviors'. *Social Networks*, 39: 84–97. DOI: <https://doi.org/10.1016/j.socnet.2014.06.001>
- Neal, ZP.** 2022. 'backbone: An R package to extract network backbones'. *PLOS ONE*, 17(5): e0269137. DOI: <https://doi.org/10.1371/journal.pone.0269137>
- Östborn, P and Gerding, H.** 2014. 'Network analysis of archaeological data: a systematic approach'. *Journal of Archaeological Science*, 46: 75–88. DOI: <https://doi.org/10.1016/j.jas.2014.03.015>
- Peebles, MA.** 2019. 'Finding a place for networks in archaeology'. *Journal of Archaeological Research*, 27: 451–499. DOI: <https://doi.org/10.1007/s10814-019-09127-8>
- Peebles, MA and Bischoff, RJ.** 2023. 'Archaeological networks, community detection, and critical scales of interaction in the U.S. Southwest/Mexican Northwest'. *Journal of Anthropological Archaeology*, 70: 101511. DOI: <https://doi.org/10.1016/j.jaa.2023.101511>
- Peebles, MA and Mills, BJ.** 2018. 'Frontiers of marginality and mediation in the U.S. Southwest: A social networks perspective'. In: Harry, K and Herr, S (eds.) *Life Beyond the Boundaries: Constructing Identity in Edge Regions of the North American Southwest*. Boulder: University Press of Colorado, pp. 25–56. DOI: <https://doi.org/10.5876/9781607326960.c002>
- Peebles, MA and Roberts, JM, Jr.** 2013. 'To Binarize or Not to Binarize: Relational Data and the Construction of Archaeological Networks'. *Journal of Archaeological Science*, 40(7): 3001–3010. DOI: <https://doi.org/10.1016/j.jas.2013.03.014>
- Prignano, L, Morer, I and Diaz-Guilera, A.** 2017. 'Wiring the Past: A Network Science Perspective on the Challenge of Archaeological Similarity Networks'. *Frontiers in Digital Humanities*, 4: 13. DOI: <https://doi.org/10.3389/fdigh.2017.00013>
- Riris, P and Oliver, J.** 2019. 'Patterns of style, diversity, and similarity in Middle Orinoco rock art assemblages'. *Arts*, 8(2): 48. DOI: <https://doi.org/10.3390/arts8020048>
- Serrano, MÁ, Boguñá, M and Vespignani, A.** 2009. 'Extracting the multiscale backbone of complex weighted networks'. *Proceedings of the National Academy of Sciences*, 106(16): 6483–6488. DOI: <https://doi.org/10.1073/pnas.0808904106>
- Tumminello, M, Micciché, S, Lillo, F, Piilo, J and Mantegna, RN.** 2011. 'Statistically validated networks in bipartite complex systems'. *PLoS ONE*, 6(3): e17994. DOI: <https://doi.org/10.1371/journal.pone.0017994>
- Visser, RM.** 2021. 'Dendrochronological Provenance Patterns. Network Analysis of Tree-Ring Material Reveals Spatial and Economic Relations of Roman Timber in the Continental North-Western Provinces'. *Journal of Computer Applications in Archaeology*, 4(1): 230–253. DOI: <https://doi.org/10.5334/jcaa.79>
- Wilkinson, TJ.** 2003. *Archaeological Landscapes of the Near East*. Tucson: University of Arizona Press. DOI: <https://doi.org/10.2307/j.ctv1jf2ddx>
- Yassin, A, Cherifi, H, Seba, H and Togni, O.** 2025. 'Exploring weighted network backbone extraction: A comparative analysis of structural techniques'. *PLOS ONE*, 20(5): e0322298. DOI: <https://doi.org/10.1371/journal.pone.0322298>
- Yener, KA.** (ed.) 2005. *The Amuq Valley Regional Projects 1: Surveys in the Plain of Antioch and Orontes Delta, Turkey, 1995–2002*. Chicago: Oriental Institute of the University of Chicago.

TO CITE THIS ARTICLE:

Nincheri, D and Moderato, M. 2026. Backbone Extraction in Archaeological Similarity Networks: Methodology and a Ceramic Case Study from the Amuq Valley. *Journal of Computer Applications in Archaeology*, 9(1): 312–328. DOI: <https://doi.org/10.5334/jcaa.281>

Submitted: 11 March 2026 **Accepted:** 10 September 2026 **Published:** 22 September 2026

COPYRIGHT:

© 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <https://creativecommons.org/licenses/by/4.0/>.

Journal of Computer Applications in Archaeology is a peer-reviewed open access journal published by Ubiquity Press.