Transport networks and towns in Roman and early medieval England: an application of PageRank to archaeological questions

Stuart Brookes and Hoai Nguyen Huynh

Abstract

This paper examines the development of a road network through time to consider its relationship to processes of urbanisation in Roman and early medieval England. Using a popular network measure called PageRank, we classify the importance of nodes in the transport network of roads and navigable waterways to assess the relative location of urban places. Applying this measure we show that there is a strong correlation between the status of towns in both Roman and medieval periods and their proximity to transport nodes with high values of PageRank. Comparison between two temporally distinct networks—Early Roman, and that recorded in the Domesday survey of AD 1086—allows for a discussion about the determinants of urban growth at different times. The applicability of PageRank to other forms of network analysis in archaeology are offered in conclusion.

Highlights

- Ranking the importance of nodes in a transport network using PageRank
- Assessing town locations in Roman and early medieval times
- Comparing the relative importance of road and riverine transport

Keywords

Roman Britain; Early medieval towns; transport networks; PageRank; network analysis

1. Introduction

Applications of network science (Gross and Sayama, 2009; Lewis, 2009; Wasserman and Faust, 1994) to other fields have yielded interesting results and proven that promising areas of research can emerge (see for e.g. Barabasi and Oltvai, 2004; Barthelemy, 2011; Borgatti et al., 2009; Dunne et al., 2002; Emirbayer and Goodwin, 1994; Glaisyer, 2004; McPherson et al., 2001; Proulx et al., 2005; Singh, 2005; Sorenson et al., 2006; Watts and Dodds, 2007). In archaeology, there is a burgeoning literature utilising such approaches (see, for e.g. Broodbank, 2000; Brughmans, et al., 2016; Collar et al., 2015; Hage and Harary, 1991; Knappett, 2011, 2013). Notably, these have included attempts to define network properties from distributions of archaeological materials such as physical monuments (e.g. Johansen, et al. 2004), artefacts of known provenance (e.g. Sindbæk, 2007), historically-attested journeys or physical route networks (e.g. Graham 2006; Isaksen, 2008), and other cultural phenomena such as language and legal traditions (e.g. Collar, 2013; Terrell, 2010).

Relevant to this paper, are works that have applied network analysis to explore the structure of past transport networks. In an early application of this type, Dicks (1972) evaluated the Roman road system of Britain via path ordering. More recently, Orengo and Livarda (2015) have tested the same evidence to demonstrate the relation between the trading activities of towns and the connections of nearby transport links.

Extending these debates, this paper seeks to demonstrate the applicability of using a popular network measure called PageRank to archaeological questions. An algorithm used by Google Search, PageRank (Brin and Page, 1998) was originally designed as a way of measuring the relative importance of web pages based on the links among them. The intended purpose of such ranking was to filter the web pages and return the most relevant ones in response to a query given to the search engine. However, the algorithm itself is universal and can usefully be applied to other situations where one seeks to identify the importance of entities in a networked system (for example, citation networks (Bollen et al., 2007; Ma et al., 2008), biological networks (Banky et al., 2013; Fletcher and Wennekers, 2017; Ivan and Grolmusz, 2011), human movement (Gleich, 2015; Jiang et al., 2008), linguistics (Esuli and Sebastiani, 2007), or even networks of graduates' education institutes and employers (Schmidt and Chingos, 2007)). Jiang (2009), in particular, has shown that there is a strong correlation between the topological structure of human movement and PageRank scores in modern street patterns. In applying this idea, we here advance the hypothesis that the relative rank of urban sites is correlated with the scale of connectivity visible in the associated transport network. As roads fossilise past patterns of movement, it might be expected that there exists a relationship between the degree of connectivity and the status of urban settlements, where the success of high-order functions is partly dependent on their ability to concentrate the flow of information, people, and materials.

Applying this hypothesis to Roman and medieval England we show that there is a strong correlation between the status of towns and their proximity to ranked transport nodes. This analysis also shows that temporal factors in the use and development of the transport network are at play. Of particular interest are distinctions between the use of the transport network of roads and navigable waterways at different times. The applicability of PageRank to other forms of network analysis in archaeology is offered in conclusion.

2. Data: towns and transport links in Roman and early medieval England

In general terms, the extents of the road and urban network of Roman Britain are well understood. Roman roads have been the subject of considerable research, notably by Thomas Codrington (1918), Ivan D Margary (1973), and latterly 'The Viatores' – a loose association of Roman road enthusiasts, who in collaboration with the Ordnance Survey Archaeology Section, sought to improve on Margary's general road system. While more recent initiatives by, for example the Roman Roads Research Association (RRRA, 2016), continue to improve our understanding of Roman roads in Britain, including nuancing chronological changes in the use of the network, in broad terms the main routes of Roman Britain are known, some incomplete branches notwithstanding (**Fig. 1a**).

Set alongside the evidence for roads, is that of Roman settlement. Lists of putative Roman towns, reconstructed from epigraphic, documentary, and archaeological sources, have been compiled—amongst others—by Millet (1990) and Burnham and Wacher (1990). Although the precise status of these centres can only be estimated from these fragmentary sources, some basic categories are clear. Millet (1990: 102–3) lists 22 'public towns' comprising *civitas*-capitals, *colonia* and *municipia*. Below these most important places are 95 smaller centres—

'small' towns—categorised by Burnham and Wacher (1990) as 'undefended settlements', 'minor towns', 'minor defended sites' and specialised 'religious' and 'industrial' sites (Millet 1990, 154–6). While it might be debated how useful such categories are, conflating and simplifying a variety of features visible at these places, they nevertheless form a benchmark from which to examine potential relationships with the transport infrastructure, and are therefore retained in this analysis.

While some Roman roads continued to be used through the medieval period (in some cases until today), this was not always the case. Recent work by Ann Cole (2013) and The Leverhulme Trust funded 'Travel and Communication in Anglo-Saxon England' project (2016), have gone some way towards clarifying our understanding of the route-way infrastructure of early medieval England. These works have sought to draw together physical archaeological evidence for roads, place-names documented in early sources containing descriptive labels for route-ways (e.g. Stratton/Stretton (Old English *strǣt-tūn* 'main/paved road estate')), along with other literary and historical evidence. The resulting route map (**Fig. 1b**) shows that there was considerable infilling of the Roman road network, but also in some cases, the partial abandonment of earlier routes in favour of new ones.

This new network of route-ways was largely determined by the landscape of settlement that emerged in the post-Roman period. Following a hiatus in urban dwelling during the fifth to seventh centuries AD, new towns gradually emerged over subsequent centuries, sometimes at the same locations as former Roman towns, but more often than not, at new locations (cf. e.g. Astill, 2000; Hodges, 2012). By the time of the Domesday survey—the 'Great Survey' of much of England and parts of Wales completed in 1086 by order of King William—a number of putative 'urban' places can be identified. (Because of the incompatibilities between Roman and Domesday evidence, this analysis restricts itself only to England.) Domesday Book associates some 112 settlements with 'burgesses' or townsmen, but of these, perhaps only 72 are likely to have been 'proper' towns at the time of the Norman Conquest (i.e. given full entries in Domesday Book; containing more than 49 'burgesses'). The relative rank of these places can further be gauged both from the textual depiction of places in Domesday Book (eight are given whole sections at the beginning of each county circuit; 35 further places have their own entries, usually prefixed by italic 'IN' and are written in capitals in a different colour; etc.), and the cumulative value of assets listed within their respective entries.

A final dataset used in the analysis of both Roman and Domesday towns, is that of navigable waterways. Accounts of the maximum extent of high medieval (*c.* AD 1300) river navigation have been compiled by a number of authors (e.g. Caffyn, 2010; Edwards, 1987; Edwards and Hindle, 1991; Langdon, 1993, 2007), which have recently been digitized by Eljas Oksanen (forthcoming). With no comparable textual sources for Roman and early medieval times, our analysis assumes that riverine navigation did not differ between the periods (but, on this point, see further discussion by Baker and Brookes, 2013: 172–4; Blair, 2007: 12–13; on Roman navigation see also Jones, 2012).

These foregoing datasets were combined in the analysis as point (centroids of town locations) and line (transport routes or navigable waterways) geometry. Because of the existence of a number of dangling ends amongst the route-way data, where the precise course of roads cannot be reliably established, cleaning operations were carried out to establish a connected graph of roads. Nodes were systematically added to all junctions between roads, or roads and waterways (**Fig. 2a**); where roads did not intersect precisely but passed within 500m of each other, bridging lines were added, along with nodal points where necessary (**Fig. 2b**). Finally, the network was simplified by removing extraneous nodes, producing an undirected graph consisting only of intersections and end points, but maintaining the length of connecting links (**Fig. 2c**). Sometimes, a road or a group of roads can be quite distant (more than 1km away) from the largest cluster of connected roads (called the largest connected component of the road network, in which there is at least a continuous path between all pairs of its nodes), and hence, are not included in the analysis. In terms of the number of nodes, less than 10% of the total number of nodes are excluded from the analysis for being unconnected.

It is noted, of course, that our understanding of route-ways, particularly in the early medieval period, but also in Roman times, is patchy, partly because of the differential survival of evidence, and partly due to more intense research carried out in some regions. The route-way datasets used in this analysis reflect our current state of knowledge, with the caveat that better understanding of Roman and early medieval transport networks may lead to different results.

3. Applying PageRank analysis

3.1 PageRank—a method of link analysis

A network typically comprises many nodes and the links among them that determine their relationship. A common task is to quantify the connections between nodes and identify the most important ones. This topic is of great interest to network science (i.e. the assessment of the importance of nodes in a network), because it can be applied to all systems that can be represented as networks. Such systems occur in many contexts ranging from human relations (e.g. network of people's interactions (Wasserman and Faust, 1994)), biology (e.g. network of cells within human body (Barabasi and Oltvai, 2004)), ecology (e.g. food chains (Dunne et al., 2002)), to technology (e.g. power systems, computer communication (Pastor-Satorras and Vespignani, 2004)) and engineering (e.g. infrasture systems, transportation (Barthelemy, 2011)). Identifying key, important entities in these systems enables one to target the right ones for practical purposes.

Network analysis offers a wide range of methods to analyse the properties of nodes and links. Among them, a number of centrality measures have been developed to characterise the (relative) importance of nodes (and, in some cases, links). Different centrality measures were developed for different purposes, focusing on different properties of the network. The most commonly used centrality measures are betweenness, closeness, degree and eigenvector (Wasserman and Faust, 1994). In the following, each measure is described and its suitability for transport network analysis is discussed. To keep the discussion simple, we assume that the network is undirected, meaning that if there is a link between a pair of nodes a and b, one can either go from a to b or vice versa from b to a. In general, this is, of course not always the case when the link between a and b might represent some unreciprocated relationship; for example, one-way roads where traffic is only allowed to flow in one direction or waterways whose flow is downstream.

3.1.1 Betweenness centrality

Betweenness centrality (Brandes, 2001; Freeman, 1978), as its name suggests, measures how often a node in a network lies on the shortest path between all pairs of other nodes. A node is deemed more important if one passes through it more times going from one part of the network to the other. Mathematically, the betweenness centrality of a node n in network $\mathcal{N}$ is defined as

$$\mathfrak{b}(n) = \sum_{a,b \in \mathcal{N}} \frac{\sigma(a,b \mid n)}{\sigma(a,b)}, \quad (1)$$

in which $\sigma(a,b)$ means the number of shortest paths between the nodes a and b, and $\sigma(a,b \mid n)$ the number of such paths passing through n. The summation is performed over all pairs of node (a,b) in $\mathcal{N}$.

There is a related measure for links, in which a link is perceived as more important if it is traversed more times as part of the shortest paths in the network. Similarly to Eq. (1), the betweenness centrality of a link e is given by

$$\mathfrak{b}(e) = \sum_{a,b \in \mathcal{N}} \frac{\sigma(a,b \mid e)}{\sigma(a,b)}, \quad (2)$$

which is similar to betweenness centrality defined in Eq. (1) except that $\sigma(a,b \mid e)$ now refers to the number of shortest paths between the nodes a and b that pass through edge e.

The idea of betweenness centrality works very well in identifying the key nodes and links in the transport network of major roads, especially ones near the centre of the network (**Fig. 3**), but it naturally underestimates the importance of nodes in the outer parts of the network, as fewer nodes there do not produce many short paths to pass through them.

3.1.2 Closeness centrality

Closeness centrality (Freeman, 1978) measures how close a node is to the centre of the network. Like betweenness centrality, the idea is based on the shortest path in the network, but instead of considering shortest paths between all pairs of nodes in the network, closeness centrality considers the shortest paths from one node to all other nodes in the network, and takes the reciprocal of sum of length of all such paths as the measure value. Formally, the closeness centrality of a node n is given by

$$\mathfrak{c}(n) = \frac{N-1}{\sum_{a=1}^{N-1} d(a,n)}, \quad (3)$$

in which N is the number of nodes in $\mathcal{N}$, and $d(a,n)$ the length of shortest path *between* nodes a and n. (To be very precise, this should be the distance *from* a *to* n, and $d(a,n)$ is generally not the same as $d(n,a)$ if the network is directed, but since we focus here only on an undirected network, such a distinction is unnecessary.) As described, the measure is only suitable for identifying nodes near to the centre of the network (**Fig. 3**).

3.1.3 Degree centrality

Degree centrality measures the importance of a node as the fraction of other nodes in the network that the current node is connected to. That means the more connections a node has, the higher is its value of degree centrality. Mathematically, the degree centrality of a note n is given by

$$\mathfrak{d}(n) = \frac{k(n)}{N}, \quad (4)$$

in which $k(n)$ means the number of nodes that are connected to n. Intuitively, this idea works well in ideal cases but more often than not, it overestimates a node's importance if the node has misleading connections such as roads leading to dead ends in the case of spatial networks.

3.1.4 Eigenvector centrality

All three centrality measures discussed above were originally developed in the context of social networks (Freeman, 1978). When applied to spatial networks, such as transportation, they suffer from either underestimating the importance of nodes in the outer parts of the network (betweenness and closeness centrality) or overestimating the importance of nodes having poor links (degree centrality). Improvements in measuring the importance of nodes in a network can be made by considering both the topology of the network and the quality of the links that each node has. One way is to examine the mutual effect of a link between two nodes on their own importance in a recursive manner, i.e. a node's importance is dependent on its neighbours' importance, which in turn is dependent on its own importance. This takes on board that not every link contributes equally to the importance of a node but rather benefits from the connection it has with an important node. The importance of each node is determined by finding the solution to a 'balanced' set of equations that describe the mutual contribution between a pair of connected nodes,

$$x(n) = \frac{1}{\lambda} \sum_{m \in B(n)} x(m), \tag{5}$$

in which $B(n)$ is the set of neighbours to which node n is connected, and λ is some constant. Mathematically, this results in a so-called eigenvector problem (Pinski and Narin, 1976; Saaty, 1977). The centrality measure arising from this is called eigenvector centrality (Newman, 2010). There are many different variants of eigenvector centrality measures that have been developed in the literature (see for e.g. Katz, 1953; Newman, 2010). In this work, we chose to employ the PageRank centrality (Brin and Page, 1998), which has proved to be successful in measuring the importance of nodes based on their connections in a network.

3.1.5 PageRank

Let us denote the to-be-determined importance of a node n as $x(n)$ which is dependent on the connections that n has with other nodes, i.e. its neighbours. Assume that n is connected to $k(n)$ neighbours which are denoted as $v_n(i)$ $(i = 1,2,\dots,k(n))$. The contribution of a connection to a node's importance is perceived as an equal share of the important of its neighbour (for that connection) among all the neighbours of that neighbour in its turn. In other words, if the node n is connected to a node a, which has $k(a)$ neighbours, n receives $\frac{1}{k(a)}$ of a's importance, $x(a)$, for that connection. This can be expressed as

$$x(n) = \sum_{i=1}^{k(n)} \frac{x[v_n(i)]}{k[v_n(i)]}. \tag{6}$$

This equation applies for all nodes n in $\mathcal{N}$, so we have a system of N equations, one for each of the nodes in the network, which form an eigenvector problem as mentioned above. In the original development of PageRank, Brin and Page used the idea of a web-surfer who randomly follows the links between webpages to calculate the chance that a page will be visited by that surfer. From the page n where the surfer starts at, he can randomly visit any of the $k(n)$ pages that n has links to, with equal probablity $\frac{1}{k(n)}$. Therefore, the chance that a page is visited is just the sum of such probability of incoming links from all its neighbours (**Fig. 4**). It was further assumed that after some point, the surfer gets bored and decides not to follow the links among the pages but instead chooses to visit a random page selected from the entire network. This would occur for a fraction β of all the time the surfer visits the pages. When this happens, the contributions to the importance of n come from *both* its immediate neighbours *and* the random pages that the surfer visited right before landing at n. Eq. (6), therefore, has to be rewritten as

$$x(n) = \frac{1-\beta}{N} + \beta \sum_{i=1}^{k(n)} \frac{x[v_n(i)]}{k[v_n(i)]} \tag{7}$$

to account for the second contribution from all random pages in the network. The quantity $x(n)$ calculated using Eq. (7) is the PageRank centrality of a node n in the network. A feature of PageRank score is that it is inherently normalised. That means the PageRank scores of all nodes in the network sum up to unity. This can be easily seen by writing Eq. (7) for each and every node in the network and summing them up; the terms on the right-hand side will automatically cancel each other resulting in unity.

3.2 Application of PageRank method to towns and roads

PageRank can be applied to analyse the structure of a route-way network and quantify the importance of transport nodes. In a network of roads, where two nodes (junctions) are linked by a road, they influence one another in a manner similar to web pages (Sec. 3.1.5, above). In other words, a junction benefits from having a direct connection to an important junction, and all junctions connected to an important junction receive equal shares of that junction's importance. The two terms in Eq. (7) above can be regarded, respectively, as 'endogenous' and 'exogenous' contributions to the importance of a node in the transport network. The latter derive from neighbouring nodes, whereas the former refer to the intrinsic value of the node itself.

After pre-processing the data (Sect. 2., above) we calculate the PageRank score $PR(n)$ of all nodes n in the connected graph of networked route-ways. Two different versions of PageRank can be computed: weighted and unweighted. In the unweighted version, all connections are treated equally and Eq. (7) is used in the calculation. In the weighted version (Xing and Ghorbani, 2004), each link between a pair of nodes carries a weight to reflect their particular relationship. A higher weight value means the pair are closer, and hence, could influence each other more easily. This is in fact what we expect to happen in reality. For example, in a transport network, a node that is close to an important node should benefit more than one further away, in other words, the weight of a road linking two junctions decreases with its length, i.e. the actual distance tracing the physical transport link. Several methods have been proposed to weight transport links, with the inverse power-law of length is commonly used (Wei-Chien-Benny and Tzai-Hung, 2015). However, in our analysis, weighted PageRank produced similar results to the unweighted version (see Sec. 5 for a discussion); thus, for the sake of simplicity, we calculated only the unweighted PageRank scores for all nodes. The factor β (called damping factor) in Eq. (7) is set to 0.85, as suggested in the literature (Brin and Page, 1998).

Once the PageRank scores have been calculated, the nodes are ranked in descending order, i.e. the node n with highest score $PR(n)$ has rank $R_n = 1$ and the node n' with lowest score $PR(n')$ has rank $R_{n'} = N$, where N is the total number of nodes in the network. It is reasonable to assume that a town's importance is reflected by the importance of transport nodes surrounding it (based on the nodes' ranking using PageRank score). Let us consider a town T, and identify the transport nodes i that are within a distance of $5km$ from T (see Sec. 4.1). The highest PageRank score among all these nodes is assigned to T, and the relative importance of town T is given by the normalised rank r_m of that node with highest local PageRank score,

$$r_m = \frac{R_m}{N}.$$

The smaller r_m is, the more important the town T is, in terms of transport links (**Fig. 5**).

The normalised rank r_m can be conveniently presented as a percentage to give us the impression of the relative importance of a town near to it. In the analysis, we consider the top 15% (rounded to nearest integer), i.e. $r_i \leq 0.15$, transport nodes to be important.

4. Route-way networks and town location in Roman and early medieval England

4.1 Roman towns

The Romanisation of Britain involved the establishment of urban sites of various kinds along with the development of a transport infrastructure. It is therefore axiomatic that the spatial pattern of Roman roads relates in some way to that of Roman towns. Using the PageRank calculation based on connections of transport nodes, it can be demonstrated that more important towns were strategically located near to well-connected transport nodes; 'near' in this analysis being defined as within 5km – approximately 1 hour's walking distance (Ohler 1989: 101) (**Fig. 6**). Of 117 Roman towns, 58 (49.6%) were located within 5km of transport nodes in the top 15% of highest PageRank score (**Table 1**); including 19 places in England identified by Millet as 'public towns' (perhaps significantly, the two 'public towns' of Millet's list that lie in Wales—Caerwent and Carmarthen—are not located near high rank transport nodes when Wales is included in the analysis; the only English exception—*Isurium*—was near a node ranked 15.7%). Two additional places deemed by Mattingley (2006: 268–9) to be significant urban sites (Corbridge and Water Newton), also fit this description. Of the remaining 38 'small' towns, a large proportion (22/57.9%) had a military function/origin as a fort, slightly above the proportions of all 'small' towns that were forts (i.e. 45.3%)('origin' as defined by Millet, 1990: 154–5) . Similarly, a marginally greater number (16/42.1%) originated as Late pre-Roman Iron Age settlements (37.9% of all 'small' towns). In both cases, this might support the view that nodal centrality developed relatively early in the evolution of the transport network, as might be expected if roads were laid out as part of the process by which a new territory was occupied and pacified. Perhaps paradoxically, 'small' towns that originated as putative 'communications' hubs at road intersections, typically display a low PageRank score: only two of 17 such sites are within the top 15% of transport links. It might be argued that such sites were a secondary development to that of the transport network, rather than determining it.

Amongst the 'small' towns, six of the eight settlements defined by Burnham and Wacher (1990) as 'industrial' sites, are close to nodes of highest 15% PageRank score. Also, near these transport nodes are all but one (*Magnis* (Kenchester, Herefordshire)) of those sites identified by Burnham and Wacher as 'potential cities' or 'possible *civitas* capitals' but ranked by Millet amongst 'small' towns. By contrast, most of those places categorised by Burnham and Wacher as 'undefended settlements', 'minor towns', 'minor defended sites' and specialised 'religious' settlements were poorly connected to nodes of high PageRank.

Addition of riverine connections to the analysis changes the relative ranking of nodes, particularly along the eastern seaboard of England, where major rivers such as the Trent, Nene, Great Ouse, and Thames provide significant inland navigations. Their effect on the connectivity on Roman towns is, however, relatively limited. Only 35 of the 117 Roman towns (29.9%) see an improvement in their location relative to transport nodes when navigable waterways are included. However, amongst the 58 most connected places, it is striking that the nodal location of six 'public' towns was significantly improved by the addition of waterways, with two (*Lindum* (Lincoln) and *Isurium* (Aldborough)) seemingly owing their importance to a nodal position on terrestrial and riverine routes. Indeed, the Trent/Ouse and Severn catchments appear to have been particularly significant in determining the nodal centrality of settlements, a finding that accords well with the distribution of imported exotica into Britain, particularly during the Middle and Late Roman phases (Orengo and Livarda 2015). By way of contrast, because it is already well-served by roads, *Londinium's* connectivity improves only slightly if the River Thames is included in the analysis. In yet other cases, riverine movement appears to have taken advantage away from 'small' towns such as Brampton and *Derventio* (Little Chester, Derbyshire).

4.2 Urban development and transport networks through time: Roman and Domesday networks compared

It is commonly assumed that Roman roads formed the backbone of later transport developments in England; indeed, many maps purportedly showing early medieval route-ways in fact show Roman roads (e.g. Hill, 1981; Pelteret, 1985). Certainly, many Roman roads, or at least their alignments, did carry on into the Middle Ages, but their survival depended on at least two basic premises: whether they were fit-for-purpose, and whether they led to where people wanted to go. In order to explore these questions, PageRank analysis was applied to the reconstructed early medieval route network and compared to the locations of towns recorded in Domesday Book. A key question this analysis sought to address was whether an examination of the transport infrastructure helps to understand, firstly, why certain Roman towns were re-established in the medieval period and others weren't, and secondly, whether there was a correlation with the longer-term success or failure of these places?

While civic and religious associations undoubtedly were the greatest determinants for urban renewal (cf. e.g. Blair, 2005: 246–90; Carver, 2010: 127–45), transport utility may also have played a part. Of the 112 Roman towns in England listed by Millet (1990), only 23 had been re-established by the late eleventh century; but perhaps significantly 20 of these were settlements that had been located within 5km of Roman transport nodes in the top 15% of highest PageRank score (**Table 2**; **Fig. 7**). Whilst the other three re-founded sites were not located near important transport nodes in Roman times, all are better connected to the early medieval network. These sites are, therefore, potentially significant in showing the development of the transport infrastructure over the first millennium AD. In the case of Worcester, the relatively poorly connected Roman 'industrial' town had by Domesday become a major shire town straddling a transport node with a top 1% PageRank score (when waterways are included). Here, it would seem that the emergence of a central place drove the construction of transport links connecting it to the rest of the network. In other cases, pre-existing transport connections influenced the development of towns. Roman *Vindonium* (Neatham, Hampshire), although an important early medieval royal and hundredal manor on the former Roman road from Chichester to Silchester, was poorly connected to transport nodes, so it is not surprising that by the 12th or 13th century the local focus of business had moved 2km away to Alton, on the main road between Winchester and London (Page, 2005).

Of the 20 Roman 'public' towns in England, 13 had been re-established as towns by the time of the Domesday survey, a further two by AD 1150. All eight settlements that on Domesday evidence might be convincingly understood as 'cities', had their origins as Roman 'public' towns that were located near very well connected former Roman transport nodes, especially with rivers. Indeed, among these cities, it can be observed that Exeter, Gloucester and Lincoln all benefited hugely from nearby navigable waterways. Amongst those that were not re-established were *Isurium Brigantum* (Aldborough) which was not located close to an important Roman road node, and two sites that had become significantly less well connected by early medieval times, supporting the idea that the infilling of the Roman road network paralleled the emergence of new nodal locations. In the case of Caistor-by-Norwich, a shift in settlement location to nearby Norwich saw a reorganisation of route-ways serving the settlement. Likewise, an attempt to re-found Roman *Petuaria* (Brough-on-Humber) in the early 13th century (Alison, 1979: 95), may have been ultimately unsuccessful because by this date there existed no transport nodes of note in its vicinity. Amongst the highest status Roman towns that failed to re-emerge were *Viriconium* (Wroxeter) and *Calleva Atrebatum* (Silchester), but in these cases an explanation from the perspective of transport geography cannot be put forward.

Comparison between the 15% nodes of highest PageRank score in Roman and early medieval times show that there was a subtle shift in connectivity across the whole network (**Fig. 8**). During the Roman period areas such as Essex and the Somerset coast were particularly well served by route-ways, perhaps in part reflecting main vectors of Continental connections via the Severn and Thames estuaries. By Domesday, new areas of dense connectivity are visible particularly in Worcestershire, where good evidence has allowed for the detailed reconstruction of a large number of saltways (Hooke, 1985: 124–6), and around the fringes of the Wash, a zone of very high levels of early medieval activity (Blair, forthcoming).

In this and other ways the Domesday urban landscape differed from that of Roman Britain. PageRank analysis show that 39 of 72 (54.2%) towns were located within 5km of transport nodes in the top 15% of highest PageRank score, an increase from Roman times (**Table 3**). Such a finding might be expected if we consider that by 1086 urban and network development had evolved together for a millennium. However, unlike Roman towns, early medieval centres appear to have been preferentially sited to access maritime connections. 29 of 72 (40.3%) see an improved location relative to transport nodes when navigable waterways are included. In the case of major cities such as Lincoln and York, connection to river transport appears to have been especially important, as they are not otherwise close to important road nodes. The PageRank method also highlights the importance of new nodes in the southwest of the country at the intersections of roads and waterways.

While the majority of Domesday towns were located within 5km of transport nodes in the top 15% of highest PageRank score, some important places were relatively poorly connected, most notably in the midlands. If one considers, for example, only the main county towns that by the eleventh century were central-places for the administration of the midland shires, it is surprising to find that at least a third were located well away from major transport hubs (**Table 4**). As many of these settlements appear to have been elevated in status during the tenth century as part of the extension of the shire system into the midlands, it is possible that routeways connecting these places had yet to fully develop.

Further indications that the evolution of the route-way network and the success of towns over longer timescales is suggested also by the analysis of those 40 settlements that have 'burgesses' named in Domesday Book, but might in other ways not classify as 'proper' towns (**Table 5**). 67.5% (27) of these places are not located within 5km of transport nodes in the top 15% of highest PageRank score; 85% (34) when only roads are considered. Amongst these poorly connected places are 12 that were no longer towns by AD 1300, including the major Domesday borough of Torksey (Hadley and Richards 2016), which, though well connected by river, was poorly served by nearby major road nodes.

5. Discussion: PageRank and the analysis of networks

5.1 Choice of PageRank over other centrality measures

The foregoing analysis suggests that PageRank is particularly well suited to quantifying the relative importance of nodes in transport networks. The seeming correspondence between the PageRank of transport nodes and the relative status of the nearby towns suggests that there is some causal link between urban status and transport connectivity. By comparison, three other commonly applied measures of centrality, namely betweenness, closeness and degree centrality, either under- or over-estimate the importance of nodes in a number of scenarios (**Fig. 3**). These biases are manifest when applied to the same road and river data. For example, transport nodes in the southwest and southeast of England would not rate highly if rankings were based on betweeness centrality, because nearby transport nodes are rarely on the shortest paths in the transport network. PageRank centrality, by contrast, does assign these nodes greater importance because these places are linked to other important nodes in the network. It should be noted that degree centrality is known to produce similar results to PageRank (Perra and Fortunato, 2008). This is also reflected in Fig. 3 where the transport nodes with highest degree centrality produce comparable pattern to those with highest PageRank scores. However, degree centrality gives more emphasis to junctions in midland England, ignoring the evident importance of coastal locations. Alternatively, degree centrality over-emphasises the importance of nodes with dangling connections – in effect, increasing the significance of places where the partiality of evidence allows us to reconstruct only part of the course of a route-way. It has been reported that degree centrality measure and PageRank produce identical results only when all the nodes have the same degree (Grolmusz, 2015), which is not the case in this work.

It is noteworthy that the idea of nodes having greater importance due to connections with other important nodes in the network is a common feature of PageRank and eigenvector centrality described in Sec. 3.1.4. Eigenvector centrality has been recently applied to archaeological problems (Collar *et al.*, 2015; Gjesfjeld, 2015; Golitko and Feinman, 2015); however, it has been noted that eigenvector centrality, as defined in Eq. (5), can be unstable and produce significantly different results when there are slight changes made to the network (Costenbader and Valente, 2003; Mills *et al.*, 2013; Gjesfjeld, 2015). Indeed, this was also observed in our analysis. PageRank, on the other hand, appears to be more stable, perhaps due to the damping term included in Eq. (7). Furthermore, the mathematical form of PageRank provides a more sensible interpretation of the flows between nodes in a transport network, where not only endogenous and exogenous values are taken into account, but also how contributions from each node are distributed to its neighbours (i.e. node n contributing $\frac{x(n)}{k(n)}$ of its importance to its neighbours depending on the number of neighbours in PageRank calculation, rather than a flat contribution $\frac{x(n)}{\lambda}$ that is globally apportioned in eigenvector centrality measure).

5.2 Weighted PageRank

As briefly discussed in Sec. 3.2, adjustment can be made to the implementation of the PageRank calculation. One important consideration is the relative weighting of connections in the network. It can be assumed that the weight of a link between two nodes must be related to the physical length of that link. In the PageRank algorithm (Langville and Meyer, 2005; Page et al., 1999), the weight of a link represents the strength of connection between two nodes. The larger the weight is, the stronger that connection, and therefore, the more influence one node has on the other. The greater the distance between two nodes—i.e. the length of the link—the less weight that connection carries in the network, and the less benefit nodes can draw from each other (Xing and Ghorbani, 2004).

In our analysis, the weighted PageRank scores do not seem to provide a better match between the calculated importance of a transport node and its nearby town(s), as compared to the unweighted PageRank scores. There could be many reasons for this. First of all, it is observed that the length of road segments connecting the junctions are approximately uniform. As a result, the weight, which is based on such length, gives little difference. In cases where the road lengths are not equal, the effect of weighting will come in to play. However, we do not really know how the physical length of a road would affect the actual relationship between the two junctions that it connects, apart from that being a decreasing relation. Therefore, *a priori*, we feel that applying equal weights to all connections, i.e. unweighted calculation, is a fair treatment.

5.3 General observations

Political and military factors were likely to have been major drivers of urban foundation during the Roman Empire, particularly in the early stages of Roman rule. PageRank analysis of the contemporary transport network, suggests that road connections laid out between these centres were especially important, providing for the mobilisation of military assets and the efficient administration of the province. The role of navigable waterways in connecting places appears to have been less important on the basis of the PageRank results. Partly, this observation is likely to reflect temporal changes in the Roman economy. Orengo and Livarda (2015, 33) suggest that the distribution of exotic goods into Roman Britain appears not to have been carried out via rivers after the Early Roman period, presumably after the establishment of the full road network used in this analysis. Likewise, the proliferation of 'small' towns in the Late Roman period may account for the poor correlation between classes of 'undefended settlements', 'minor towns', 'minor defended sites' and specialised 'religious' settlements with nodes of high PageRank, as these came to the fore as a secondary development to the transport network. Future work could usefully examine chronological variation in the spatial patterning of such settlements.

If Late Roman towns can be characterised as centres of consumption, early medieval towns by contrast were more orientated towards production (cf. Palliser, 2000, 21–4), with an attendant different relationship to the transport network. PageRank analysis suggests that the use of navigable rivers (and by implication the sea) became much more significant in determining the location and status of towns. In part, this may reflect something of the quality of early medieval roads and the preference for riverine over terrestrial movement (cf. Stenton, 1936); but the frictionless transport of commodities is also likely to have been determining. Shipping was an important aspect of the medieval economy and the ranking of nodes connecting rivers and roads reflects this orientation.

Significantly, the PageRank analysis provides some possible insights into the success and failure of different towns over the longue durée. In both Roman and early medieval examples there is a good correlation between the relative status of towns, defined in terms of legal status or fiscal value, and their proximity to transport nodes of high connectivity. And in both cases it can be shown that settlements that were poorly connected were unlikely either to re-emerge or flourish as central-places. Of the Roman 'public' towns that failed to become early medieval towns, most were—or became—less well connected by 1086. Likewise, 12 of the 15 places listed with burgesses in Domesday Book but failed as towns over the course of the later Middle Ages, were also ones located away from major transport hubs. This close correspondence between the long-term development of a town and its proximity to important transport nodes, in some cases (e.g. Worcester) driving the development of a transport infrastructure, reinforces the view that transport connectivity was at least one precondition for urban success. While factors other than the PageRank of nearby transport nodes were also important in determining the success of urban places over time, this study suggests that an examination of network connectivity provides useful data for further analysis.

5. Conclusion

With network analyses becoming more frequently applied methods in archaeology it is important to appreciate the relative merits of different forms of link analysis. In this study PageRank is shown to be a powerful addition to the cannon – producing results that can be used to make broader observations about the dynamics of urban development in England during the first millennium AD. While PageRank has been used in other contexts (see Sec. 1) it has yet to be used widely in archaeology. Two examples to date have both sought to use the technique to mine archaeological datasets, either to estimate the archaeological potential of urban deposits (Dubbini and Gattiglia 2014), or predictively model historic landuse (Dubbini and Lodoen 2016). Here, we have argued that PageRank has potential in archaeological applications examining physical networks. Other applications of this approach are manifold, including riverine and maritime networks, trade links, itineraries, and urban street patterns, as are analyses of cultural, social and linguistic connections, wherever these can be presented in the form of a network.

Acknowledgements

HNH acknowledges the support of A*STAR International Fellowship. 'Travel and Communication in Anglo-Saxon England' is a major Research Project funded by the Leverhulme Trust (Ref: RPG-2014-074). Data used in this analysis was generously provided by Eljas Oksanen, the National Monuments Record, Museum of London Archaeology, Clwyd-Powys Archaeological Trust. We are also grateful to the following for commenting on early versions of this paper: Andrew Bevan, Alex Langlands, Alessio Palmisano, Eleanor Rye, Barbara Yorke.

Figure captions

Figure 1. (a) Roads and towns in Roman Britain; (b) Early medieval route-ways and principal settlements of England listed in Domesday Book.

Figure 2. A description of the cleaning operations carried out to produce a connected graph of route-ways. (a) Two roads A and B with segments defined by the corresponding sequence of points (A_1, A_2, A_3, A_4) and (B_1, B_2, B_3, B_4) intersect in the middle of their segments. When the point of intersection is not recorded in either A or B, a new point I was added to both of them, resulting in the new sequences (A_1, A_2, I, A_3, A_4) and (B_1, B_2, I, B_3, B_4), respectively. (b) Two non-intersecting roads A and B with segments defined by the corresponding sequence of points

(A_1, A_2, A_3, A_4) and (B_1, B_2, B_3, B_4) are connected if they are within 500m of each other. New points are added if necessary as in 2a. (c) Network of points on roads A and B with segments defined by the corresponding sequence of points (A_1, A_2, I, A_3, A_4) and (B_1, B_2, I, B_3, B_4) is simplified to a graph of only intersections and end points by eliminating intermediate nodes of degree $k = 2$, keeping the end points of degree $k = 1$ and intersections of degree $k > 2$. The length of the remaining connections are preserved, (i.e. the same as when the intermediate points are present.

Figure 3. Comparison of identification of top 15% most important transport nodes in the network of Roman roads using four different centrality measures, namely betweenness centrality, closeness centrality, degree centrality and PageRank. This illustrates the suitability of PageRank over the other three commonly used measures in network analysis when the identified top nodes show good correlation with important Roman towns.

Figure 4. Illustration of idea behind PageRank. The values associated with the links indicate the amount of importance that the target nodes earn from those connections. For example, the importance of node A is the sum of all the contributions from its neighbours $\frac{x(B)}{k(B)}, \frac{x(C)}{k(C)}, \frac{x(D)}{k(D)}, \frac{x(E)}{k(E)}$ and $\frac{x(F)}{k(F)}$, and a fixed term as described in Eq. (7).

Figure 5. An illustration of connectivity score of a town. Only the three nodes N_2, N_3 and N_8 (crosshatched) within 5km distance of the town T (hollow) are considered. The connectivity score of T is given by $\max\{PR(N_2), PR(N_3), PR(N_8)\}$.

Figure 6. Locations of Roman towns and nearby top transport nodes.

Figure 7. Locations of Domesday towns and nearby top transport nodes.

Figure 8. Kernel density estimation plots of the top 15% of transport nodes in the Roman (left) and early medieval (right) networks. Also depicted are rankings of the top 15% of nodes.

References

Alison, K.J., 1979. Victoria County History, Yorkshire (East Riding), vol. iv. Victoria County History, London.

Astill, G., 2000. General survey 600–1300, in: Palliser, D. (Ed.) The Cambridge Urban History of Britain. Cambridge University Press, Cambridge, pp. 27–49.

Baker, J. and Brookes, S., 2013. Beyond the Burghal Hidage. Brill, Leiden.

Banky, D., Ivan, G., and Grolmusz, V., 2013, Equal Opportunity for Low-Degree Network Nodes: A PageRank-Based Method for Protein Target Identification in Metabolic Graphs, PLOS ONE 8, 1-7.

Barabasi, A.-L. and Oltvai, Z. N., 2004, Network biology: understanding the cell's functional organization, Nat. Rev. Genet. 5.2, 101-113.

Barthelemy, M., 2011, Spatial networks, Phys. Rep. 499.1-3, 1-101.

Blair, J., 2005. The Church in Anglo-Saxon Society. Oxford University Press, Oxford.

Blair, J., 2007. Introduction, in: Blair, J. (Ed.) Waterways and Canal-building in Medieval England. Oxford University Press, Oxford, pp. 1–18.

Blair, J., forthcoming. Building Anglo-Saxon England. Princeton University Press.

Bollen, J., Rodriguez, M. A., and Van de Sompel, H., 2007, MESUR: Usage-based Metrics of Scholarly Impact, in Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries JCDL '07, 474

Borgatti, S. P., Mehra, A., Brass, D. J., and Labianca, G., 2009, Network Analysis in the Social Sciences, Science 323.5916, 892-895.

Brandes, U., 2001, A faster algorithm for betweenness centrality, Journal of Mathematical Sociology 25.2, 163-177.

Brin, S. and Page, L., 1998, The anatomy of a large-scale hypertextual Web search engine. Computer Networks ISDN 30, 107–17.

Broodbank, C., 2000. An Island Archaeology of the Early Cyclades. Cambridge University Press, Cambridge.

Brughmans, T., Collar, A., and Coward, F., 2016, The Connected Past. Oxford University Press, Oxford.

Burnham, B.C. and Wacher, J., 1990. The 'Small Towns' of Roman Britain. Batsford, London.

Caffyn, D.J.M., 2010. River Transport 1189–1600. Unpublished PhD thesis: University of Sussex

Carver, M.O.H., 2010. The Birth of a Borough. An archaeological study of Anglo-Saxon Stafford. The Boydell Press, Woodbridge.

Codrington, T., 1918. Roman Roads in Britain, 3rd ed. Society for Promoting Christian Knowledge, London.

Collar, A., 2013. Re-thinking Jewish Ethnicity through Social Network Analysis, in: Knappett, C. (Ed.) Network Analysis in Archaeology: New Approaches to Regional Interaction. Oxford University Press, Oxford, pp. 223–246.

Collar, A., Coward, F., Brughmans, T. and Mills, B. J., 2015, Networks in Archaeology: Phenomena, Abstraction, Representation, J Arch. Method and Theory 22.1, 1–32.

Costenbader, E., and Valente, T. W. 2003, The stability of centrality measures when networks are sampled, Social Networks 25.4, 283–307.

Dicks, T.R.B., 1972, Network Analysis and Historical Geography, Area 4.1, 4–9.

Dubbini, N. and Gattiglia, G., 2013, A PageRank based predictive model for the estimation of the archaeological potential of an urban area, 2013 Digital Heritage International Congress (DigitalHeritage), Marseille 2013, 571—578.

Dubbini, N. and Lodoen, A. 2016. Statistical and mathematical models for archaeological data mining: a comparison. https://www.academia.edu/7748525/Statistical_and_mathematical_models_for_archaeological_data_mining_a_comparison (accessed 16/06/17)

Dunne, J. A., William, R. J., and Martinez, N. D., 2002, Network structure and biodiversity loss in food webs: robustness increases with connectance, Ecology Letters 5.4, 558-567.

Edwards, J.F., 1987. The Transport System of Medieval England and Wales – a geographical synthesis. Unpublished PhD thesis: University of Salford.

Edwards, J.F. and B.P. Hindle, 1991. The Transportation System of Medieval England and Wales. J. of Hist. Geog. 17(2), 123–34.

Emirbayer, M. and Goodwin, J., 1994, Network analysis, culture, and the problem of agency, American Journal of Sociology 99.6, 1411-1454.

Esuli, A. and Sebastiani, F., 2007, PageRanking WordNet synsets: An Application to Opinion-Related Properties, In Proceedings of the 35th Meeting of the Association for Computational Linguistics, Prague, Czech Republic, 424–431.

Fletcher, J. M. and Wennekers, T., 2017, From Structure to Activity: Using Centrality Measures to Predict Neuronal Activity, International J. of Neural Systems, 1750013

Freeman, L.C., 1978. Centrality in social networks conceptual clarification. Soc. Networks 1.3, 215–239.

Gjesfjeld, E., 2015. Social Network Analysis of Archaeological Data from Hunter-Gatherers: Methodological Problems and Potential Solutions. J. of Arch. Method and Theory 22.1, 182–205.

Glaisyer, N., 2004, Networking: Trade and exchange in the eighteenth-century British empire, Historical Journal 47.2, 451–476.

Gleich, D. F., 2015, PageRank Beyond the Web, SIAM Review 57.3, 321-363.

Golitko, M., and Feinman, G. M. 2015, Procurement and distribution of Pre-Hispanic Mesoamerican obsidian 900 BC – AD 1520: a social network analysis. J. of Arch. Method and Theory 22.1, 206–247.

Graham, S., 2006. Networks, agent-based models and the Antonine itineraries: implications for Roman archaeology. J. of Mediterranean Arch. 19, 45–64.

Grolmusz, V., 2015, A note on the PageRank of undirected graphs, Information Processing Letters 115.6-8, 633-634.

Gross, T. and Sayama, H., 2009. Adaptive Networks: Theory, Models and Applications. Springer

Hadley, D. and Richards, J., 2016. The Winter Camp of the Viking Great Army, AD 872–3, Torksey, Lincolnshire. Antiq. J. 96, 23–67.

Hage, P. and Harary, F., 1991. Exchange in Oceania: a Graph Theoretic Analysis. Clarendon Press, Oxford.

Hill, D., 1981. An Atlas of Anglo-Saxon England. Basil Blackwell, Oxford.

Hodges, R., 2012. Dark Age Economics, a new audit. Bloomsbury, London.

Hooke, D., 1985. The Anglo-Saxon Landscape: the Kingdom of the Hwicce. Manchester University Press, Manchester.

Isaksen, L., 2008. The application of network analysis to ancient transport geography: a case study of Roman Baetica. Digital Medievalist 4. http://digitalmedievalist.org/journal/4/isaksen/ (accessed 12/05/17)

Ivan, G. and Grolmusz, V., 2011, When the Web meets the cell: using personalized PageRank for analyzing protein interaction networks, Bioinformatics 27.3, 405.

Jiang, B., 2009, Ranking spaces for predicting human movement in an urban environment, International Journal of Geographical Information Science 23.7, 823-837.

Jiang, B., Zhao, S., and Yin, J., 2008, Self-organized natural roads for predicting traffic flow: a sensitivity study, Journal of Statistical Mechanics: Theory and Experiment 2008.7, P07008.

Johansen, K.L., Laursen, S.T., and Holst, M.K., 2004. Spatial patterns of social organization in the Early Bronze Age of South Scandinavia. J. of Anth. Arch. 23(1), 33–55.

Jones, J.E., 2012. The Maritime Landscape of Roman Britain. BAR British Series 556. British Archaeological Reports, Oxford.

Katz, L., 1953, A new status index derived from Sociometric Analysis, Psychometrika, 39-43.

Knappett, C., (ed.) 2011. An Archaeology of Interaction: Network Perspectives on Material Culture and Society. Oxford University Press, Oxford.

Knappett, C., (ed.) 2013. Network Analysis in Archaeology: New Approaches to Regional Interaction. Oxford University Press, Oxford.

Langdon, J., 1993. Inland Water Transport in Medieval England. J. of Hist. Geog. 19(1), 1–11.

Langdon, J., 2000. Inland Water Transport in Medieval England - the View from the Mills: a Response to Jones. J. of Hist. Geog. 26(1), 75–82.

Langville, A.N. and Meyer, C.D., 2005. A Survey of Eigenvector Methods for Web Information Retrieval, Siam Review 47.1, 135–61

Lewis, T. G., 2009. Network Science: Theory and Applications. Wiley

Ma, N., Guan, J., and Zhao, Y., 2008, Bringing PageRank to the citation analysis, Information Processing & Management 44.2, 800-810.

McPherson, M., Smith-Lovin, L., and Cook, J. M., 2001, Birds of a feather: Homophily in social networks, Annual Review of Sociology 27, 415-444.

Margary, I.D., 1973. Roman Roads in Britain, 3rd ed. J. Baker, London.

Mattingley, D., 2006. An Imperial Possession. Penguin, London.

Millet, M., 1990. The Romanization of Britain: an essay in archaeological interpretation. Cambridge University Press, Cambridge.

Mills, B., Roberts, J. M., Jr., Clark, J. J., Haas, W. R., Jr., Huntley, D., Peeples, M. A., Borck, L., 2013, The dynamics of social networks in the Late Prehispanic US Southwest, in: Knappett, C. (ed.) Network analysis in archaeology: new approaches to regional interaction. Oxford University Press, Oxford, pp. 181–202.

NetworkX library in Python, url = https://networkx.readthedocs.io/en/stable/reference/generated/%networkx.algorithms.link_analysis.pagerank_alg.pagerank.html

Newman, M., 2010. Networks: An Introduction. Oxford University Press, USA.

Ohler, N., 1989. The Medieval Traveller. Boydell Press, Woodbridge.

Oksanen, E., forthcoming. Navigable rivers in England and Wales, c. 1050–1350. Archaeology Data Service

Orengo, H.A. and Livarda, A., 2015. The seeds of commerce: A network analysis-based approach to the Romano-British transport system, J. Arch. Sci. 66, 21–35.

Page, L., Brin, S., Motwani, R. and Winograd, T., 1999. The PageRank Citation Ranking: Bringing Order to the Web

Page, M., 2005. Medieval Alton: the origins of a market town, Proc. of the Hampshire Field Club and Arch. Soc. 60, 170–4.

Palliser , D. 2000. The origins of British towns, in: Palliser, D. (ed.) The Cambridge Urban History of Britain. Cambridge University Press, Cambridge, pp. 17–26.

Pastor-Satorras, R. and Vespignani, A., 2004. Evolution and Structure of the Internet: A statistical physics approach. Cambridge University Press, Cambridge.

Pelteret, D., 1985. The Roads of Anglo-Saxon England. Wiltshire Arch. and Nat. Hist. Magazine 79, 155–63.

Perra, N. and Fortunato, S., 2008, Spectral centrality measures in complex networks, Phys. Rev. E 78, 036107.

Pinski, G. and Narin, F., 1976, Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics, Information Processing & Management 12.5, 297-312.

Proulx, S. R., Promislow, D. E. L., and Phillips, P. C., 2005, Network thinking in ecology and evolution, Trends in Ecology and Evolution 20.6, 345-353.

RRRA (Roman Roads Research Association), 2016, http://www.romanroads.org/ (accessed 31.05.2017)

Saaty, T. L., 1977, A scaling method for priorities in hierarchical structures, Journal of Mathematical Psychology, 15.3, 234-281.

Schmidt, B. M. and Chingos, M. M., 2007, Ranking Doctoral Programs by Placement: A New Method, PS: Political Science & Politics 40.3, 523-529.

Sindbæk, S.M., 2007. The small world of the Vikings. Networks in Early Medieval communication and exchange. Norwegian Arch. review 40(1), 59–74.

Singh, J., 2005, Collaborative networks as determinants of knowledge diffusion patterns, Management Science 51.5, 756-770.

Sorenson, O., Rivkin, J. W., and Fleming, L., 2006, Complexity, networks and knowledge flow, Research Policy 35.7, 994-1017.

Stenton, F.M., 1936. The road system of medieval England. Econ. Hist. Review 7 (1), 1–21.

Terrell, J.E., 2010. Language and material culture on the Sepik Coast of Papua New Guinea: using Social Network Analysis to simulate, graph, identify, and analyze social and cultural boundaries between communities. J. of Island and Coastal Arch. 5(1), 3–32.

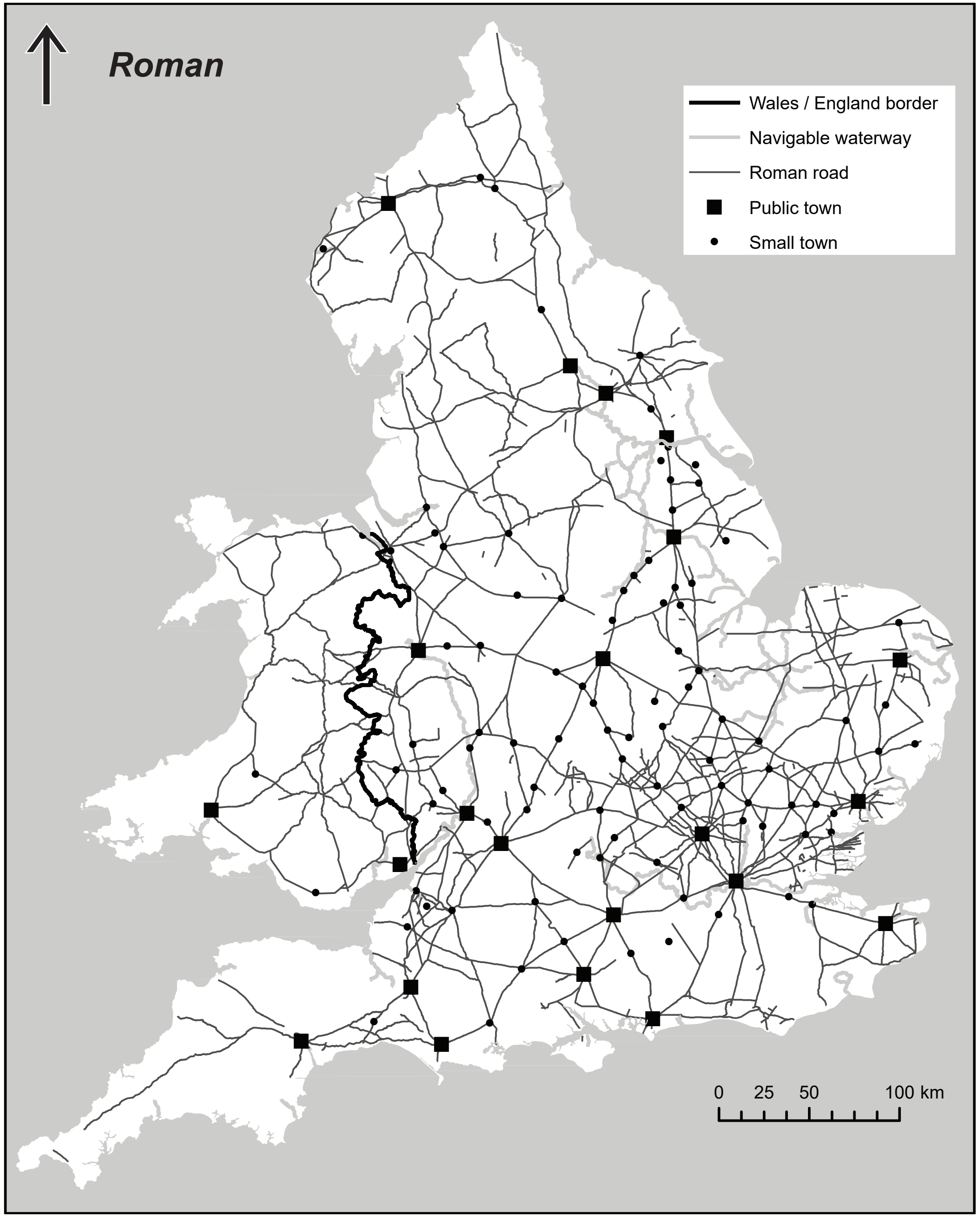
Travel and Communication in Anglo-Saxon England, 2016, http://www.ucl.ac.uk/archaeology/research/directory/travel-communication-anglo-saxon-england (accessed 31.05.2017)

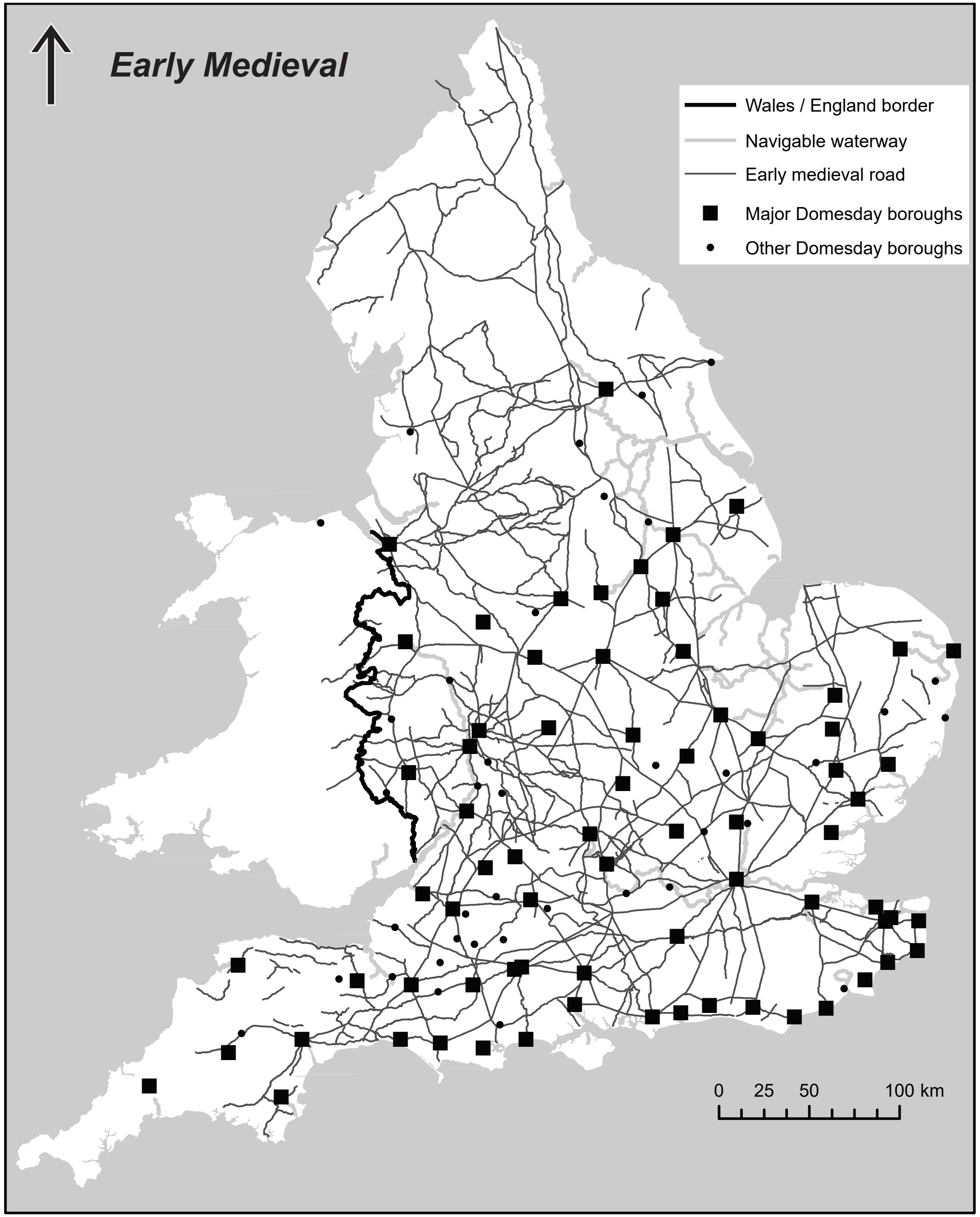
Wasserman, S. and Faust, K., 1994. Social Network Analysis: Methods and Applications: Cambridge University Press, Cambridge

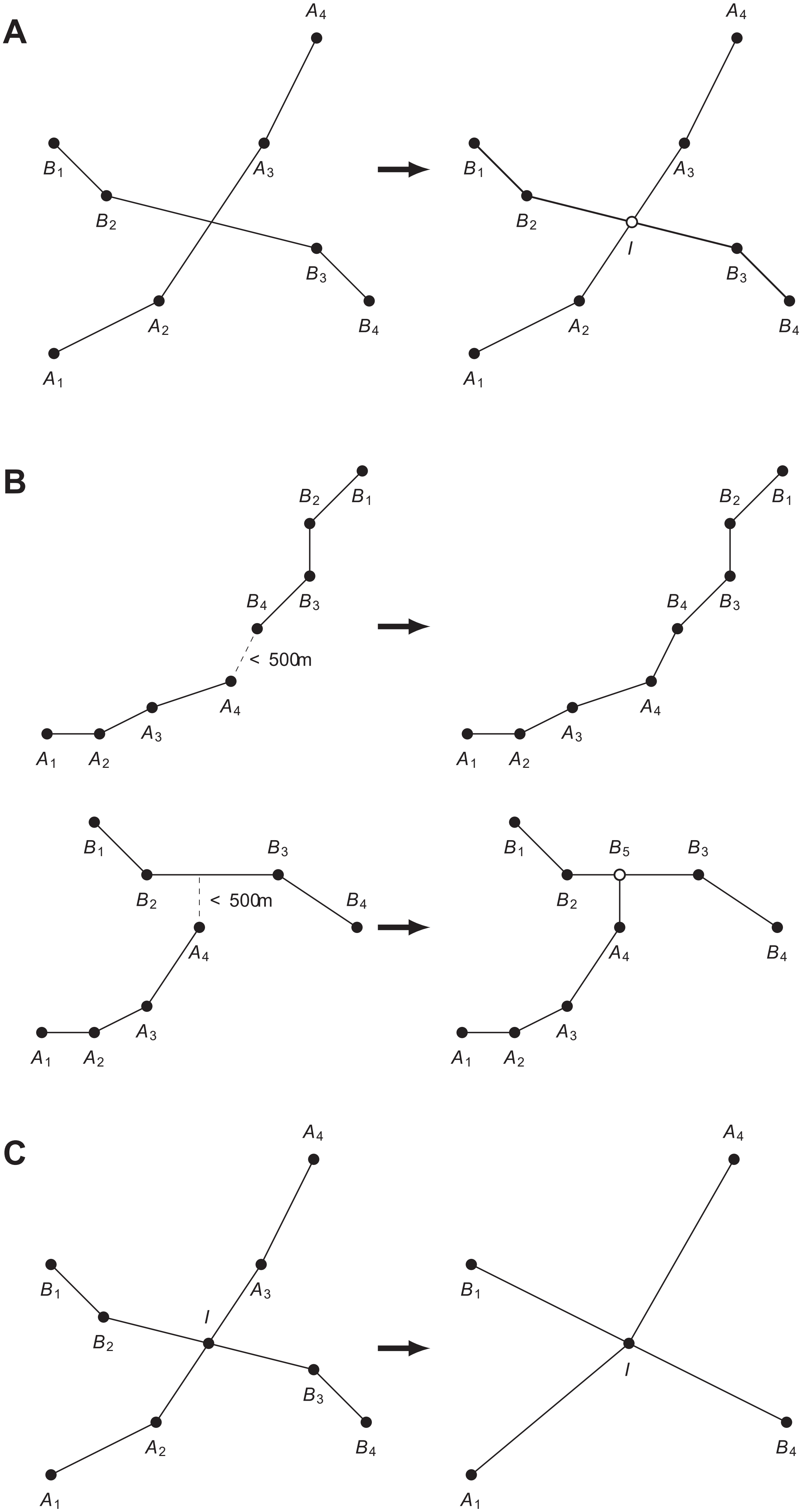
Watts, D. J. and Dodds, P. S., 2007, Influentials, Networks, and Public Opinion Formation, Journal of Consumer Research 34.4, 441.

Wei-Chien-Benny, C. and Tzai-Hung, W., 2015. Geographically Modified PageRank Algorithms: Identifying the Spatial Concentration of Human Movement in a Geospatial Network, PLOS ONE 10, 1–23.

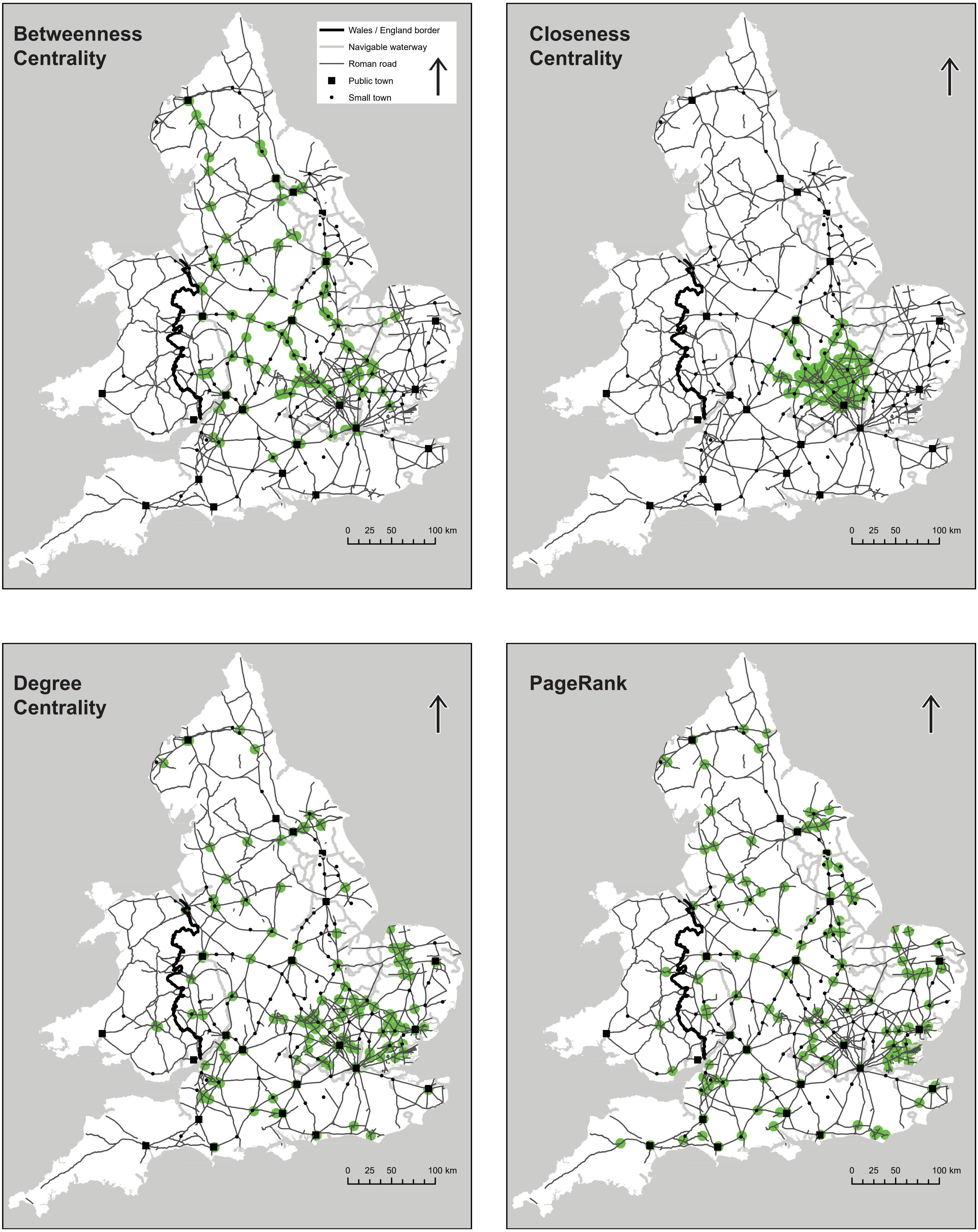
Xing, W. and Ghorbani, A., 2004, Weighted PageRank algorithm, Proceedings of Second Annual Conference on Communication Networks and Services Research 2004, 305-314.


Roman
Wales / England border
Navigable waterway
Roman road
Public town
Small town
0 25 50 100 km

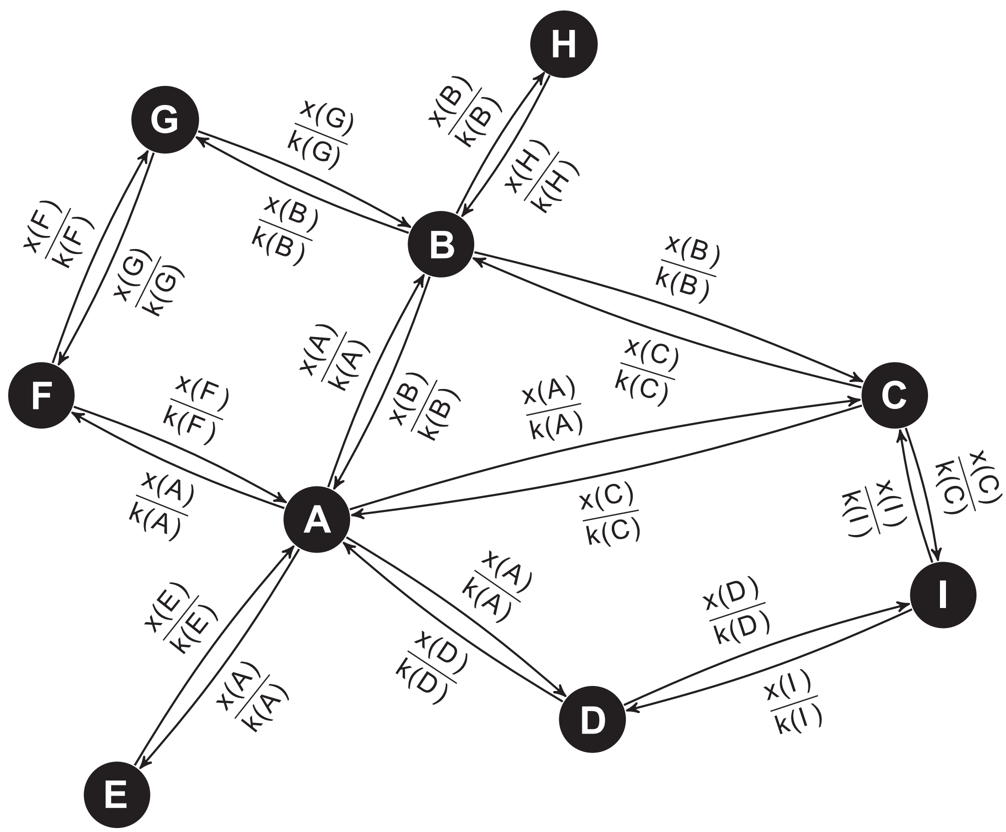

Early Medieval
Wales / England border
Navigable waterway
Early medieval road
Major Domesday boroughs
Other Domesday boroughs
0 25 50 100 km

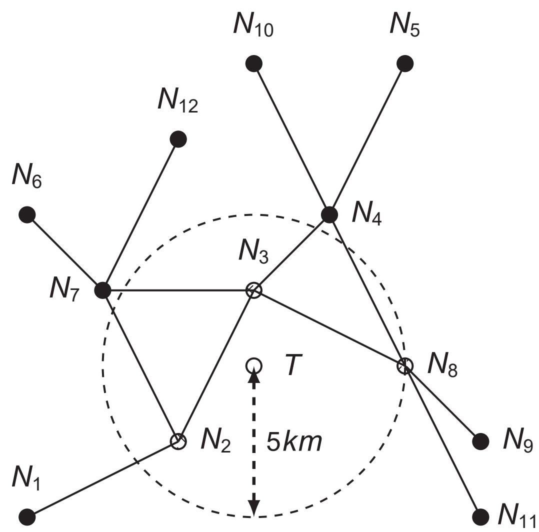


A
A_4
B_1
A_3
B_2
I
B_3
A_2
B_4
A_1
B
B_2
B_1
B_4
B_3
< 500m
A_4
A_3
A_1
A_2
B_5
C

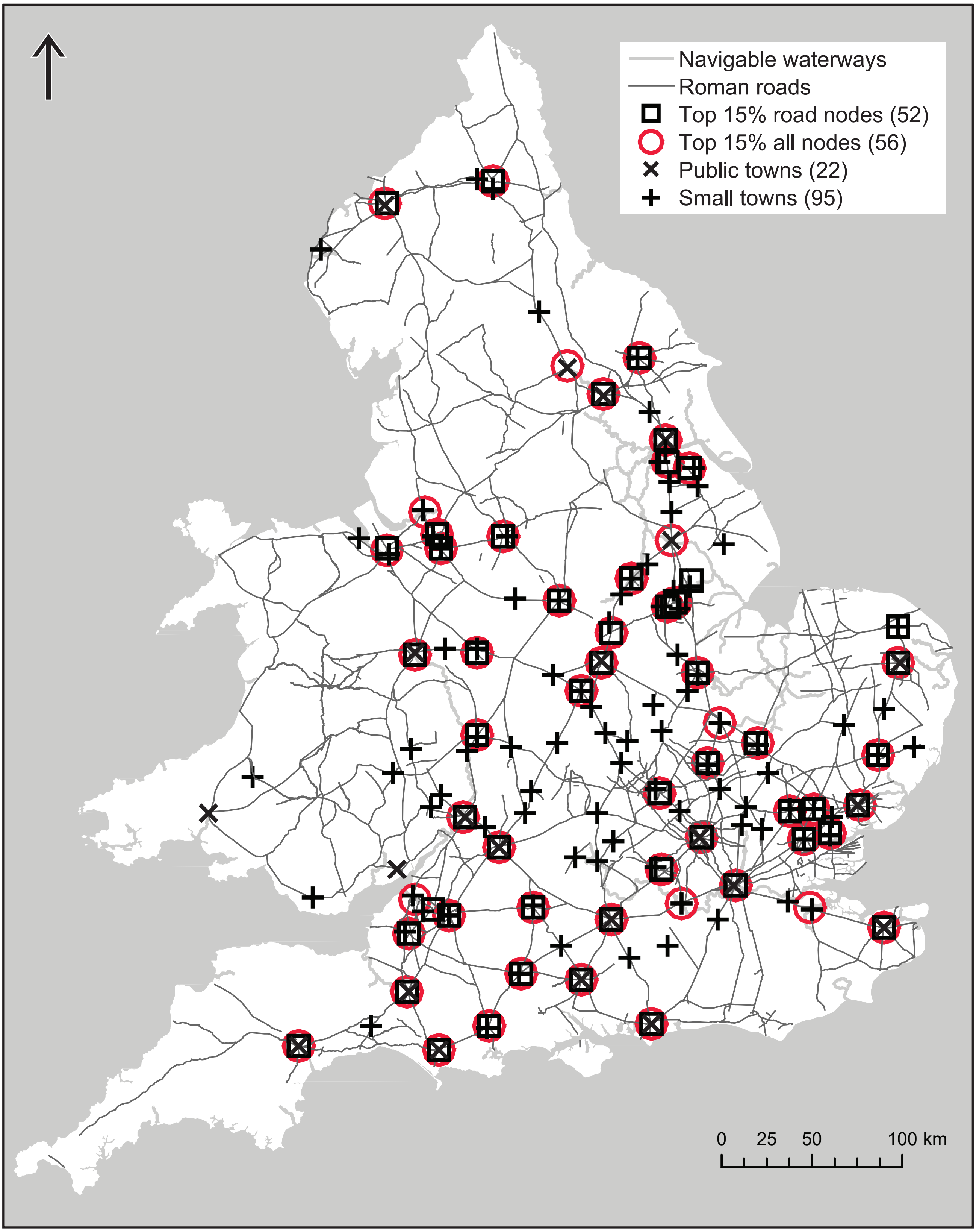


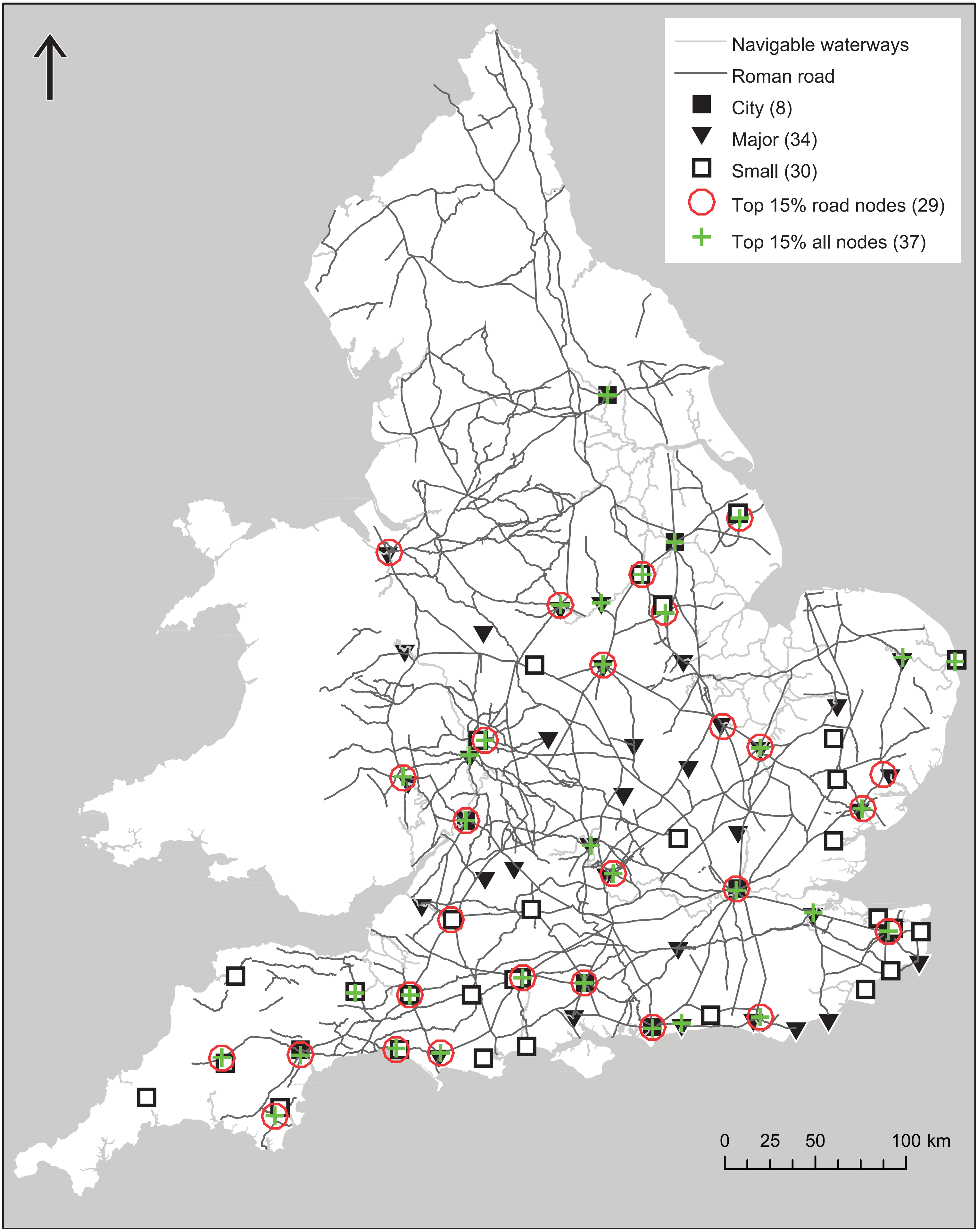
Betweenness Centrality
Wales / England border
Navigable waterway
Roman road
Public town
Small town
0 25 50 100 km
Closeness Centrality
0 25 50 100 km
Degree Centrality
0 25 50 100 km
PageRank
0 25 50 100 km


H
G
B
F
C
A
I
D
E
$\frac{x(G)}{k(G)}$
$\frac{x(B)}{k(B)}$
$\frac{x(B)}{k(B)}$
$\frac{x(H)}{k(H)}$
$\frac{x(F)}{k(F)}$
$\frac{x(G)}{k(G)}$
$\frac{x(B)}{k(B)}$
$\frac{x(C)}{k(C)}$
$\frac{x(A)}{k(A)}$
$\frac{x(B)}{k(B)}$
$\frac{x(F)}{k(F)}$
$\frac{x(A)}{k(A)}$
$\frac{x(A)}{k(A)}$
$\frac{x(C)}{k(C)}$
$\frac{x(C)}{k(C)}$
$\frac{x(I)}{k(I)}$
$\frac{x(E)}{k(E)}$
$\frac{x(A)}{k(A)}$
$\frac{x(A)}{k(A)}$
$\frac{x(D)}{k(D)}$
$\frac{x(D)}{k(D)}$
$\frac{x(I)}{k(I)}$

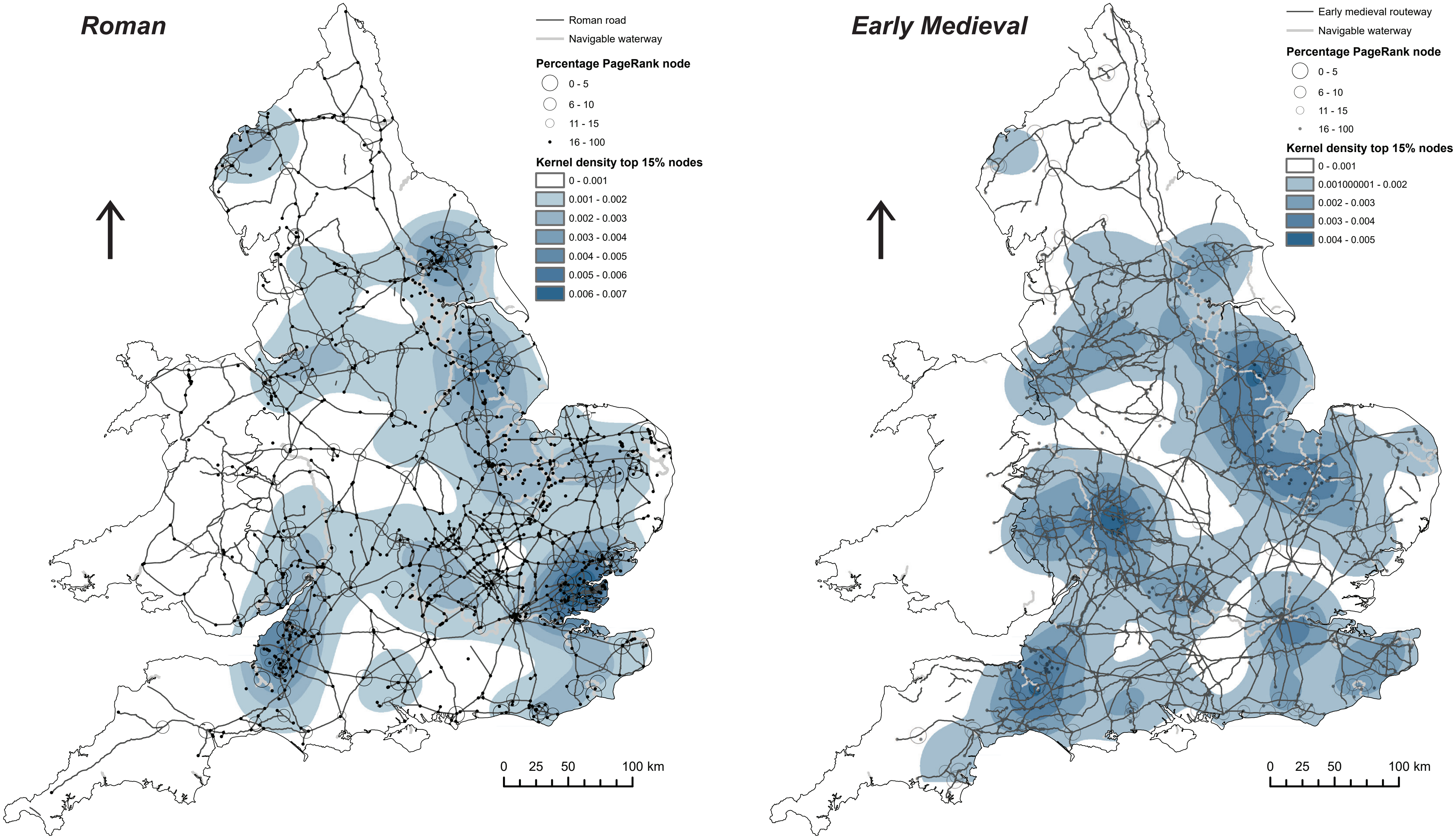


N10
N5
N12
N6
N4
N3
N7
T
N8
5km
N2
N9
N1
N11


Navigable waterways
Roman roads
Top 15% road nodes (52)
Top 15% all nodes (56)
Public towns (22)
Small towns (95)
0
25
50
100 km



Navigable waterways
Roman road
City (8)
Major (34)
Small (30)
Top 15% road nodes (29)
Top 15% all nodes (37)
0 25 50 100 km



Roman
Roman road
Navigable waterway
Percentage PageRank node
0 - 5
6 - 10
11 - 15
16 - 100
Kernel density top 15% nodes
0 - 0.001
0.001 - 0.002
0.002 - 0.003
0.003 - 0.004
0.004 - 0.005
0.005 - 0.006
0.006 - 0.007
0 25 50 100 km
Early Medieval
Early medieval routeway
Navigable waterway
Percentage PageRank node
0 - 5
6 - 10
11 - 15
16 - 100
Kernel density top 15% nodes
0 - 0.001
0.001000001 - 0.002
0.002 - 0.003
0.003 - 0.004
0.004 - 0.005
0 25 50 100 km

Table captions

Table 1. Roman towns in England (after Millet 1990, 154–5) located within 5km of the top 15% of transport nodes (value is bracketed if not in top 15%) defined by the road network (R), and the road and navigable waterway network (R+). 'Public' towns are in **Bold**.
Table 2. Roman towns in England that re-emerged as towns by the late 11th century, and their proximity to PageRank nodes of the Roman and early medieval (DB) transport networks (value is bracketed if not in top 15%).
Table 3. Domesday towns located within 5km of early medieval transport nodes in the top 15% of highest PageRank score, without (DB R) and with rivers (DB R+). Also shown is the relationship of these places to Roman transport nodes (Rom R; Rom R+). (Value is bracketed if not in top 15%.)
Table 4. Domesday county towns of Midland England, and their proximity to early medieval transport nodes (value is bracketed if not in top 15%).
Table 5. 40 lesser boroughs listed in Domesday Book, and the PageRank of transport nodes within 5km. Also shown is whether these settlements were still classed as boroughs in AD 1300.

Table 1:

Milletno	Romantown	Roman name	Description	%R	%R+
5	**Brough-on-Humber**	*Petuaria*	civitas capital	5	5
32	**Wroxeter**		civitas capital	4	0
37	**Leicester**	*Ratae Coritanorum*	civitas capital	10	10
41	**Caistor-by-Norwich**	*Venta Icenorum*	civitas capital	1	2
87	**Cirencester**	*Corinium Dobunnorum*	civitas capital	1	1
95	**Silchester**	*Calleva Atrebatum*	civitas capital	4	5
103	**Ilchester**	*Lindinis*	civitas capital	8	8
104	**Exeter**	*Isca Dumnoniorum*	civitas capital	8	0
107	**Winchester**	*Venta Belgarum*	civitas capital	2	2
109	**Chichester**	*Noviomagus Regnorum*	civitas capital	1	1
110	**Canterbury**	*Durovernum Cantiacorum*	civitas capital	0	0
111	**Dorchester**	*Durnovaria*	civitas capital	6	7
113	**Carlisle**	*Luguvalium*	civitas capital	3	4
3	**Aldborough**	*Isurium Brigantum*	civitas captial	43	16
4	**York**	*Colonia Eboracensium*	colonia; municipium	5	0
18	**Lincoln**	*Colonia Domitiana Lindensiumlindum*	colonia; municipium	21	3
67	**Gloucester**	*Colonia Nervia Glevensiumcolonia*	colonia; municipium	14	1
78	**Colchester**	*Colonia Claudia Victricensis*	colonia; municipium	6	3
81	**Verulamium**	*Verulamium*	colonia; municipium	1	1
90	**London**	*Londinium*	colonia; municipium	1	1
83	Chelmsford	*Caesaromagus*	civitas captial; definite fort	5	5
29	Saltersford	Perhaps *Causennis*	communications?	8	9
31	Willoughby	*Vernemetum*	communications?	5	7

14	Northwich	*Condate*	definite fort	15	15
23	Whitchurch	*Mediolanum*	definite fort	12	26
34	Penkridge	*Pennocrucium*	definite fort	10	10
72	Dropshot	*Magiovinium*	definite fort	13	13
7	Dragonby		definite LPRIA site	5	6
28	Sleaford		definite LPRIA site	4	19
84	Heybridge		definite LPRIA site	4	5
8	Kirmington		dual LPRIA/fort	8	8
60	Coddenham	*Combretovium*	dual LPRIA/fort	11	11
12	Wilderspool		industrial site	45	2
114	Brampton		industrial site	8	25
25	Little Chester	*Derventio*	industrial site; definite fort	7	12
101	Charterhouse-on-Mendip	Perhaps *Iscalis*	industrial site; definite fort	14	13
49	Droitwich	*Salinae*	industrial site; dual LPRIA/fort	12	13
17	Middlewich	*Salinae*	industrial site; probable fort	3	4
21	Thorpe-by-Newark	*Ad Pontem*	minor defended settlement; definite fort	13	2
58	Cambridge	*Duroliponte*	minor defended settlement; dual LPRIA/fort	11	1
46	Godmanchester	*Durovigutum*	minor town; definite fort	36	10
94	Mildenhall	*Cunetio*	minor town; probable fort	13	12
2	Malton	*Derventio*	possible civitas capital; definite fort	0	0
40	Water Newton	*Durobrivae*	potential city; definite fort	3	2
112	Corbridge	*Coriosopitum*	potential city; definite fort	13	14
99	Rochester	*Durobrivae*	potential city; definite LPRIA site	69	5
92	Sea Mills	*Abona*	probable fort	39	11
57	Sandy		probable LPRIA site	8	9
105	Badbury	*Vindocladia*	probable LPRIA site	12	12
106	Old Sarum	*Sorviodunum*	probable LPRIA site	4	5
13	Buxton	*Aquae Arnemetiae*	religious site; dual LPRIA/fort	13	13
100	Bath	*Aquae Sulis*	religious site; dual LPRIA/fort	7	3
68	Wycomb		religious site; probable LPRIA site	12	13
16	Heronbridge		uncertain	10	0
76	Great Dunmow		uncertain	2	3
77	Braintree		undefended settlement; definite LPRIA site	10	10
96	Staines	*Pontibus*	undefended settlement; probable fort	51	9
30	Sapperton	Perhaps *Causennis*	undefended settlement; probable LPRIA site	5	7

Table 2

Millet no.	**Roman town**	**Description**	**% Rom R**	**% Rom R+**	**% DB R**	**% DB R+**
111	**Dorchester**	Roman civitas capital; major Domesday town	6	7	5	6
103	**Ilchester**	civitas capital; small Domesday town, 107 messuages	8	8	1	2
37	**Leicester**	civitas capital; major Domesday town	10	10	2	3
104	**Exeter**	civitas capital; Domesday city	8	0	9	1

107	**Winchester**	civitas capital; Domesday city	2	2	2	3
109	**Chichester**	civitas capital; Domesday city	1	1	9	10
110	**Canterbury**	civitas capital; Domesday city	0	0	0	0
4	**York**	colonia / municipium; Domesday city	5	0	(48)	1
18	**Lincoln**	colonia / municipium; Domesday city	(21)	3	(33)	6
67	**Gloucester**	colonia / municipium; Domesday city	14	1	8	1
78	**Colchester**	colonia / municipium; major Domesday town	6	3	1	0
81	**Verulamium**	Colonia / municipium; as St Albans small Domesday town, 46 messuages	1	1	(36)	(48)
90	**London**	colonia / municipium; Domesday city	1	1	3	0
14	Northwich	definite fort; Probably a *de facto* market by 1086	15	15	14	13
49	Droitwich	industrial site / dual LPRIA/fort; small Domesday town, 89 messuages	12	13	14	15
50	Worcester	industrial site / probable LPRIA site; major Domesday town	(73)	(74)	(16)	1
108	Neatham (as Alton)	minor defended settlement / communications?; recorded 1086, *mercatum*	(69)	(69)	(25)	(27)
58	Cambridge	minor defended settlement; dual LPRIA/fort; major Domesday town	11	1	13	1
99	Rochester	potential city / definite LPRIA site; major Domesday town	(69)	5	(32)	6
106	Old Sarum	probable LPRIA site; small Domesday town, unspecified number of messuages	4	5	3	4
100	Bath	religious site; dual LPRIA/fort; small Domesday town, 20 messuages	7	3	11	15
96	Staines	undefended settlement / probable fort; 1218, *mercatum*.	(51)	9	(51)	(27)
19	Horncastle	mint 970s–1016, Commercial and administrative centre of the important soke of Horncastle from the 10th century.	(86)	(87)	(60)	(60)

Table 3:

Borough	**Domesday status**	**% Rom R**	**% Rom R+**	**% DB R**	**% DB R+**
Canterbury	City	0	0	0	0
Chichester	City	1	1	9	10
Exeter	City	8	0	9	1
Gloucester	City	14	1	8	1
Lincoln	City	(21)	3	(33)	6
London	City	1	1	3	0
Winchester	City	2	2	2	3
York	City	5	0	(48)	1
Arundel	Major	(44)	(20)	(60)	7
Cambridge	Major	11	1	13	1
Chester	Major	10	0	12	2
Colchester	Major	6	3	1	0
Derby	Major	7	12	5	7
Dorchester	Major	6	7	5	6
Hereford	Major	(18)	(19)	6	12
Huntingdon	Major	(36)	10	13	(19)
Ipswich	Major	(32)	(33)	11	(20)
Leicester	Major	10	10	2	3

Lewes	Major	(28)	(22)	6	4
Norwich	Major	1	2	(50)	3
Nottingham	Major	(95)	(68)	(63)	6
Oxford	Major	(49)	(16)	(19)	13
Rochester	Major	(69)	5	(32)	6
Wallingford	Major	(19)	(18)	1	2
Worcester	Major	(73)	(74)	(16)	1
Bath	Small	7	3	11	15
Bridport	Small	(86)	(87)	10	10
Droitwich	Small	12	13	14	15
Fordwich	Small	0	0	0	0
Grantham	Small	8	9	12	14
Ilchester	Small	8	8	1	2
Louth	Small	(95)	(95)	6	7
Lydford	Small	5	6	4	5
Newark	Small	(62)	13	9	1
Old Sarum	Small	4	5	3	4
Taunton	Small	(89)	(78)	(66)	2
Totnes	Small	(81)	(83)	6	4
Wilton	Small	4	5	3	4
Yarmouth	Small	(43)	(17)	(85)	14

Table 4:

County towns	**Shire**	**% DB R**	**% DB R+**
Bedford	Bedfordshire	(81)	(80)
Buckingham	Buckinghamshire	(36)	(36)
Cambridge	Cambridgeshire	13	1
Chester	Cheshire	12	2
Derby	Derbyshire	5	7
Gloucester	Gloucestershire	8	1
Hereford	Herefordshire	6	12
Hertford	Hertfordshire	(23)	(25)
Huntingdon	Huntingdonshire	13	(19)
Leicester	Leicestershire	2	3
Lincoln	Lincolnshire	(33)	6
Northampton	Northamptonshire	(44)	(43)
Nottingham	Nottinghamshire	(63)	6
Oxford	Oxfordshire	(19)	13
Shrewsbury	Shropshire	(68)	(65)
Stafford	Staffordshire (gone by 1086)	(94)	(95)
Stamford	Stamfordshire	(69)	(96)
Warwick	Warwickshire	(99)	(99)
Winchcombe	Winchcombeshire (gone by 1086)	(62)	(62)
Worcester	Worcestershire	(16)	1

Table 5:

Borough	**Town in 1300?**	**% DB R**	**% DB R+**

Ashwell	no	(81)	(80)
Axbridge	yes	(96)	(11)
Beccles	yes	(85)	(89)
Bedwyn	no	(62)	(61)
Bradford on Avon	yes	(50)	(50)
Bridlington	yes	(77)	(76)
Bruton	yes	(69)	(69)
Calne	yes	(59)	(58)
Clare	yes	(48)	(48)
Clifford	no	2	3
Dadsley	no	(34)	(40)
Dunwich	no	(43)	(43)
Ewias Harold	no	7	8
Eye	yes	(84)	(84)
Frome	yes	(54)	(57)
Langport	yes	(69)	2
Milborne Port	no	(72)	(73)
Milverton	no	(92)	(94)
Newport Pagnell	yes	(48)	(47)
Okehampton	yes	4	5
Penwortham	yes	1	1
Pershore	yes	(48)	(22)
Pocklington	yes	7	8
Quatford	no	(48)	14
Reading	yes	(61)	(67)
Rhuddlan	yes	(97)	(99)
Rye	yes	(92)	4
Southwark	yes	3	0
St Albans	yes	(36)	(48)
Stanstead Abbots	no	(23)	(25)
Tanshelf	yes	(53)	8
Tewkesbury	yes	(28)	3
Tilshead	no	(26)	(28)
Torksey	no	(37)	7
Tutbury	no	(94)	(95)
Warminster	yes	(60)	(61)
Wigmore	no	(16)	(17)
Wimborne Minster	yes	(28)	(32)
Winchcombe	yes	(62)	(62)
Windsor	yes	(97)	(28)