Speculations on browsing, directed searching, and linking in relation to the Bradford Distribution

ABSTRACT

Extensive literatures exist on information searching theory and techniques, as well as on the Bradford Distribution. This distribution, also known as "Bradford's Law of Scattering," tells us that information on a subject is dispersed in a characteristic and robust pattern that appears consistently across many different environments. This pattern may be expected to have important implications for information searching theory and techniques. Yet these two research literatures are rarely considered in relation to each other. It is the purpose of this article to distinguish three Bradford regions and speculate on the optimum searching techniques for each region. In the process, browsing, directed searching in databases, and the pursuit of various forms of links will all be considered. Implications of growth in size of a literature for optimal information organization and searching will also be addressed.

Introduction

Extensive research literatures exist on information search techniques and on the Bradford Distribution—but rarely are the two considered in relation to each other. (A search produced a handful of references linking

the two—Lancaster et al., 1991; Leimkuhler, 1977; White, 1981—but none of these items address the questions being considered in this paper.) The Bradford Distribution, or, Bradford's Law of Scattering, describes how information on a subject is distributed among the resources where such information may be expected to be found. The distribution will be described in more detail shortly, but the key point is that the distribution tells us that information is neither randomly scattered, nor handily concentrated in a single location. Instead, information scatters in a characteristic pattern, a pattern that should have obvious implications for how that information can most successfully and efficiently be sought.

It is the object of this paper to link the Bradford Distribution with searching techniques, and to consider what techniques might work the best for which regions of the Bradford distribution. In the process, I will also speculate on a variety of other possible implications of the connections between Bradford's Law and searching.

The Bradford Distribution

Samuel Bradford originally developed his data on the basis of studying the distribution of articles in journals in two areas, applied geophysics and lubrication. He studied the rates at which articles relevant to each subject area appeared in journals in those areas. He identified all journals that published more than a certain number of articles in the test areas per year, as well as in other ranges of descending frequency. He wrote:

> [I]f scientific journals are arranged in order of decreasing productivity of articles on a given subject, they may be divided into a nucleus of periodicals more particularly devoted to the subject and several groups or zones containing the same number of articles as the nucleus, when the numbers of periodicals in the nucleus and succeeding zones will be as 1:n:n² . . . (Bradford, 1948, p. 116).

In principle, there could be any number of zones, with the number of articles in each zone being the total number of articles divided by the number of zones. In his empirical data, however, Bradford identified just three zones, and three will be used here for simplicity's sake. Bradford found in his empirical data that the value of "n" was roughly 5. Suppose, then, that someone doing an in-depth search on a topic finds that four journals constitute the core, and contain fully one-third of all the relevant articles
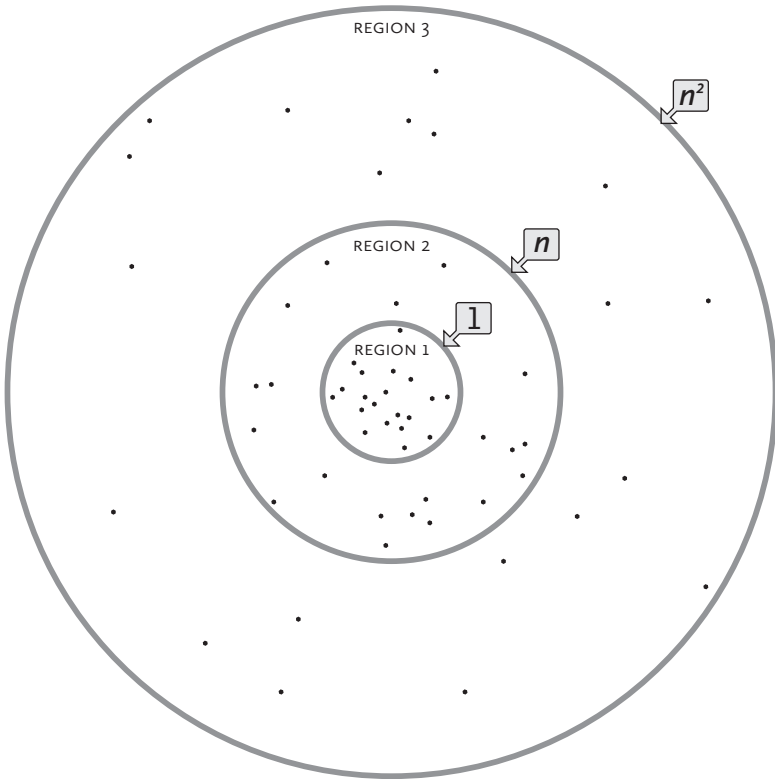
FIG. 1. *The Bradford regions. Each search region contains one-third of the articles on the subject. Each ring is five times the area of the next smaller one.*

found. ("Relevant" is used here strictly in the sense of topical relevance.) If the value for n is 5, then 4 x 5 = 20 journals will, among them, contain another third of all the relevant articles found. Finally, the last third will be the most scattered of all, being spread out over $4 \times 5^2 = 100$ journals. See Figure 1.

The distribution indicates that one could find a great many relevant articles on a topic nicely concentrated in a few core journals. But finding the rest of the relevant articles involves an increasingly more extensive search, as the average yield of articles per additional journal examined becomes smaller and smaller the farther out, i.e., the more remotely from the core journals, one goes. The Bradford data thus tell us that desired material is neither perfectly concentrated in one place, as we might hope, nor is it completely randomly distributed either, as we might fear.

There has been considerable debate about the exact formulation of the distribution. However, for our purposes, the general formulation given by Bradford himself (above) is sufficient. Many types of recorded information have been found to follow the Bradford Distribution. See reviews in Chen and Leimkuhler (1986), Fedorowicz (1982), and Wilson (1999). Recently, Hood and Wilson (2001) found the same pattern operating regarding the number of databases to search to find the literature on 14 test queries. (For simplicity's sake, the discussion in this paper is restricted to materials on a subject, though other kinds of searches, such as for materials by authors, materials coming out of research institutes or government agencies, etc., probably also follow a Bradfordian pattern.) In sum, it seems reasonable to assume, in the absence of any contrary information, that recorded information of all types on a subject, and in many different physical forms, is distributed in this same Bradfordian manner.

Searching the Bradford regions

Assuming then that information on a subject is distributed in the Bradfordian pattern, what are the implications for search strategy? The proposed relationship between the Bradford Distribution and searching techniques will first be described below in a general way, then more specific details will be developed.

From the searcher's perspective, the innermost core of the distribution—let us call it Region 1—is the richest in materials on one's subject of interest. One might therefore expect it to be the easiest to search, because there is so much wheat, so little chaff. Perhaps one might even be able to find items in Region 1 without any conventional formal search techniques at all—one may simply browse.

In the next ring, Region 2, there are still a good number of items on one's subject to be found, but they are now scattered more widely, and many non-relevant items (that is, many items not on the desired subject) are interspersed among the (topically) relevant items. In this region, it is no longer adequately efficient to search by browsing, though one can, of course, always browse if one wants. But for any reasonably efficient search effort, some kind of formal information organization of the material is needed, coupled with search techniques specifically designed to take advantage of that organization. Region 2 is where indexing and database searching come into their own. This approach is here called "directed searching." More precisely, directed searching is done in situations where some form(s) of information organization, description, and/or indexing of the

information exists, thus enabling explicit query development and search of that query.

In Region 3, the outermost area, a substantial number of relevant items are still to be found, but they are so widely dispersed that finding the occasional relevant item is rather like the proverbial problem of "finding a needle in a haystack." Here, following citation or other links from articles closer to the core may prove to be the most efficient sort of searching. Of course, links may lead to articles in any part of the distribution, from the core out to the periphery.

However, in the far edges of the distribution linking may be especially useful as a search technique. In conventional database searching terms, as one moves outward from the center, the precision gets worse and worse— one must review more and more irrelevant items to find any one relevant item, and the recall climbs more and more slowly. Here, conventional indexing/retrieval may be seen to break down. Searching through vast resources becomes inefficient, and the number of "false drops" is so high that it becomes ever harder to detect the occasional desired item. Here, in Region 3, the best solution may be to create links, or to take advantage of pre-existing links, such as article notes/citations or Web links.

Thus, at the most generalized level, we may see the core region of the Bradford distribution as the ideal location for browsing, the middle region as the ideal location for searching in the mode of indexing/retrieval, and the outer region as the ideal location for pursuing links.

Having conjectured this pattern for ideal searching, however, it will be demonstrated shortly that all three generic search techniques are currently used throughout the Bradford regions. Should they be?

Generic search techniques

We will want to examine the Bradford regions in more detail shortly, but in the meantime, it will reward us to examine these three general search techniques in more detail—browsing, directed searching, and linking.

Browsing

The first generic search technique is *browsing,* which involves successive acts of glimpsing, fixing on a target to examine visually or manually more closely, examining, then moving on to start the cycle over again. This definition is strongly based on Kwasnik's (1992) conceptualization of browsing. (See also discussion in Chang and Rice, 1993.)

We can see browsing as having several functions in information searching. First, with respect to the Bradfordian distribution of literature, *one purpose of browsing is to find the core.* Browsing, in its nature, ignores the file structure or other formal organization of information. For example, when we browse in library stacks or in a bookstore, our glimpsing of the materials on the shelves does not follow the call number or alphabetical order of the materials. In other words, we generally do not scan in a strict left to right order along the shelf. Our glance is not systematic—though we do take advantage, in a larger sense, of any relatedness in the arrangement of the materials on the shelves. Browsing is therefore a quick technique, but also a chancy one. When one does not know where the heart of a subject area is to be found, however, this rapid technique may enable one to zoom in quickly on the richest area to be searched.

A second purpose of browsing is to search the core. As the core is densely populated with relevant material, it is generally not necessary to create or negotiate one's way through additional formal information organization structures. The population of relevant items is large enough that one can simply browse the area and reliably locate a good number of desirable materials in a short time.

Two techniques described in Bates (1989) represent specific forms of this generic browsing technique. **Journal run** consists of reviewing contents pages of core journals in an area.

> [This] technique, by definition, guarantees complete
> recall within that journal, and, if the journal is cen-
> tral enough to the searcher's interests, this technique
> also has tolerably good precision (1989, p. 412).

Journal run is a clear case of browsing in the Bradford core (Region 1) of a subject. The materials are collocated in the journal because editorial decisions have brought them together, not because anyone indexed them and collocated them in an information system.

The second specific browsing technique described in Bates (1989) is *area scanning*. This technique consists of browsing the materials that are physically collocated with materials located earlier in a search. The classic example of area scanning is browsing through the classified arrangement of books on library shelves. So, in effect, area scanning is a kind of browsing that is used in the second Bradford region, where directed searching normally takes place.

Another common type of browsing is *Web-browsing*, that is, moving around via links between and within documents on the World Wide

Web. Materials from all three Bradford regions may be found on the Web. However, finding Region 3 items may be particularly successful, given the nature and structure of the Web.

Thus, some form of browsing has been used in all three regions of the Bradford Distribution. In general, the strengths of browsing are its intuitiveness (building on age-old foraging behavior of our hunter-gatherer forbears—Sandstrom, 1994) and its directness—what you see is what you get. The downside of browsing is that it is limited; we can do it only for so long and in a relatively narrow area, before we are overcome by our modern need to move on to the next task.

Directed searching

The second generic searching technique is what is here being called *directed searching,* to distinguish it from "searching" in general. Directed searching is done (and can only be done) where some sort of previous indexing, cataloging, or classification has been done on the materials being searched. To put it differently, the information is deliberately structured in order that subsets of the information can be selectively targeted for retrieval. The deliberate structuring may be as simple as the automatic coding of all individual words in a database so that they can be individually searched ("free-text" searching), or as intellectually complex as deep and detailed human indexing and cataloging.

In either case, a descriptive or metadata structure is created by human beings in front of or in addition to the original target information. The searcher then may utilize whatever searching capabilities are provided to exploit this representational structure, in order to find desired information from among a much larger set of resources.

In Bates' 1989 article, two of the techniques described there are classic directed searching techniques: **subject searches in bibliographies and abstracting and indexing (A & I) services**, and **author searching** (p. 412). Directed searching has been a primary focus of information science research and discussion over the last thirty years. Most of the 29 searching tactics described by Bates (1979) are moves that the searcher can make to promote an effective search in the context of these elaborate indexing structures and retrieval capabilities.

Directed searching has been used on all three Bradford regions. The browsable core generally is automatically included in any good search formulation, so that items located in both Regions 1 and 2 are retrieved. Region 3 items have long been found in online searching through use of extensive search formulations containing many OR'd terms ("hedges"),

as well as by searching several databases representing peripheral subject areas. The strength of directed searching is its thoroughness and extensive reach. Its weaknesses are that it lacks the directness of browsing and the targeted precision of the citation link.

The advent of the more-effective Google search engine, which is based on linkage patterns as well as conventional search engine retrieval mechanisms, suggests that heavy utilization of links in *really* large information environments, especially in the outer limits of Region 3, may be the best way to go when the Bradford territory has grown very large. On the other hand, because of the recency of the Internet, we may simply have not yet found the best forms of directed searching for it.

Incidentally, Jaime Pontigo-Martinez' dissertation work (1984) tested whether scientific experts, unaware of which Bradford ring specialist articles in their field came from, would judge the articles from all rings as being of equal value. The null hypothesis, that there would be no difference in assessed value between the rings, was supported. This result suggests that, despite the difficulties in searching for articles in outer rings, it is still valuable to locate them, if possible.

Linking

Linking long pre-dates the Internet, but has come into its own in recent years with the advent of sophisticated networking technology (Kleinberg, 1999; Lempel and Moran, 2001). *Links* are deliberate connections created by people or automatically by software between parts of a document or between documents. Linking has been utilized to connect texts with relevant related references (footnotes) for centuries. Linking may include the references from index entries to page numbers in books, the "see" and "see also" references between catalog entries, the links within and between documents on the World Wide Web, and the citation links in citation indexes, among other things.

Strictly speaking, we should distinguish between the *creation* of links by authors or Web-page designers, and the *pursuit or following* of links by searchers. It is the latter sense that is being used in association with the discussion of searching techniques in this paper.

In their nature, links may be made between records in any Bradford ring to any other ring. Thus, in principle, searching via links may move anywhere in the Bradford space. Links, while very targeted and precise, may also be capricious and incomplete. Following links, the searcher may range widely, but incompletely, across the relevant materials.

Two of these kinds of links are mentioned among the search techniques described in Bates' 1989 article, *footnote chasing*, and **citation searching** (p. 412). Footnote chasing, also called "backward chaining" (Ellis, 1989), "involves following up footnotes found in books and articles of interest, and therefore moving backward in successive leaps through reference lists" (Bates, 1989, p. 412). Each such leap, of course, takes one farther back in time.

To utilize citation searching, or "forward chaining," there must be in existence citation indexes that have already located in a single index the relationships of a large number of citing and cited works. Major citation indexes are the three produced by the Institute for Scientific Information, the *Science Citation Index,* the *Social Sciences Citation Index,* and the *Arts & Humanities Citation Index.* One searches in a citation index by looking up an earlier article or book of interest and discovering who has cited it since its publication, thus coming forward in time.

Linking has exploded with the advent of the World Wide Web; indeed the Web, even in its name, can be seen as the primary and best place to manifest links between every imaginable form of information, so long as it can be resident in a computer.

Impact of domain size

We have been skirting an important issue, which now needs to be addressed. If we think about real-world searching, the size of the domain has a very important impact on searching; specifically, different search techniques are appropriate for differently-sized bodies of information, regardless of ring. It turns out that there is a nice illustrative analogy that can be used here. It has been long recognized in the field of geography (Wilson and Bennett, 1985, p. 84) that towns and cities grow in patterns that look very much like our circular model of the Bradford Distribution. Indeed, such distributions are universal statistical phenomena. Bradford did not so much as invent his distribution, but rather applied this general phenomenon to information science.

When communities grow, they follow a certain characteristic pattern. This pattern was well illustrated in the settling of the American West. A new area opening up would draw farmers, ranchers, or miners. Then a merchant would set up a general store to serve the needs of these groups. As the community grew, the filling in of the community would follow the pattern of Figure 1, with denser settlement in the heart of town, and characteristic scattering toward the edges of town. Thus the distributional

pattern remains the same throughout the time of growth, from village to metropolis.

Nonetheless, cities do not look like villages. With the growth in size and increased density at that core, more and more kinds of activities became sustainable. Saloons, livery stables, and assay offices appeared next, followed later by the sheriff, schools, and churches. Continued growth eventually supported concert halls, libraries, operas, and many other less common institutions.

A town of regional importance can sustain more rarefied types of institutions than smaller towns in the area. So in the growth of towns, we can see that size matters a lot. Though the Bradford-style distribution characterizes the scattering of settlement from the time of the tiny beginning outpost to the large city, the kinds of institutions that these settlements will support varies tremendously with size.

The same pattern can be seen with information resources. That is, the Bradford pattern operates throughout, from the time of a tiny beginning literature known to only a handful of people to the development of a vast literature of interest to millions of people. On the other hand, the kinds of searching, the kinds of access that make sense for selecting information from this literature will vary with the size of the total domain.

Derek de Solla Price illustrated this pattern in his discussion of the early days of science. He noted that in the seventeenth century, scientists initially developed journals in order to lighten the load of reading books, personal correspondence, and other sources of information on the ever-growing scientific enterprise. Initially, the journals "had the stated function of digesting the books and doings of the learned all over Europe" (Price, 1963, p. 63). After a while, the journals themselves became important publication venues for new discoveries. Eventually, however, it became difficult to follow all the journals being published, and abstracting indexes were developed (Price, 1961, pp. 96ff). Indeed, it may well be that every major technological and intellectual development in information access in succeeding centuries has arisen out of the pressures of literature growth.

Let us now look more closely at what a Bradford-distributed literature would look like as it grows in size. When a small group of originators begins producing small numbers of documents, the absolute size of the domain is so small that while Bradford regions might be present in an incipient form, they are not yet very evident. One individual or one paper may be more important than others, and form the Region 1 focus of social and documentary networks. However, the total number of resources and connections among them are few, and interested individuals can generally find

all they want by browsing and following citations. In very small towns one does not need a map, because everything is visible on one street; likewise, one needs little or no assistance once one has come upon the "small town" of a starter literature on a subject.

With the growth in the literature, however, just as with the growth of the town, auxiliary devices are necessary to enable one to find desired locations—maps and directories for geographical locations, classifications and indexing for bibliographical regions. As the literature grows, one access method after another is shed, as more sophisticated techniques and technologies must be developed to maintain effective access to the literature.

With very small literatures, one may need only one or two distinct types of search method to do effective searching over the whole subject matter domain. In a very mature, highly developed literature, there may be half a dozen discernible Bradford regions, each, perhaps, requiring different search techniques to maximize search effectiveness.

Until the late nineteenth century, library collections, and topical literatures on given subjects were still comparatively small. As Price, Senders, and others have demonstrated, the numbers of book and journal titles have been growing exponentially since the scientific journal (seventeenth century) and the modern printed book (fifteenth century) began. In practical terms, exponential growth means that the number of titles doubles every so many years—for journals, every fifteen years (Price, 1961, p. 100), and books every 22 years (Senders, 1963, p. 1068). Such growth does not make much difference in the early days—1,000 book titles doubling to 2,000 book titles in 22 years is not particularly conspicuous—but book titles doubling from 20 million to 40 million in 22 years is *very* conspicuous! Thus the publishing and library collecting of documentary forms of various types was quite stable for centuries, with relatively glacial change. The need for innovation became particularly acute in the nineteenth century, however, as the absolute numbers of new titles quickly outdistanced every new innovation introduced.

It is probably not accidental that the modern profession of librarianship emerged at that same time. The days of the interested amateur librarian ended when the need for attention to ever-newer techniques and ever-larger buildings made library work a full-time focus of whoever took the job.

It may even be the case that the growth of collections and of materials within any given subject area reached the point in the nineteenth century that sophisticated Region 2 devices had to be introduced for the first time. Before then, catalogs were barely more than inventories of collections. With growth, however, anyone who was not a dedicated scholar in a narrow

subject field needed the selectivity offered by more advanced techniques of cataloging, indexing, and classification, in order to maneuver among the growing collections.

In the speculative spirit of this paper, I would suggest that as literatures grew over the last century and a half, each major new intellectual and physical technology for information access represented another "institution," in the geographical analogy, in the growing turf that was information. Just as larger cities can sustain orchestras and concert halls, so also can larger literatures sustain databases, bibliographies, and sophisticated indexing and online system design.

In the meantime, more sophisticated forms of browsing and linking have been introduced recently as well. Online browsing is becoming better supported through subtler and more user-centered system design. The introduction of the citation indexes in the 1960s and 1970s enabled forward chaining as a linking method, and, of course, the Internet created an explosion of linking possibilities for the searcher.

Final speculations: Best searching techniques?

We have seen in the prior discussion that on the one hand, browsing, directed searching, and linking, respectively, might best be matched with Bradfordian Regions 1, 2, and 3. On the other hand, we have seen that all three generic techniques have been used with all three regions, indeed, have generally been used in the absence of thinking about the Bradford Distribution altogether. A profusion of technologies continues to develop, which supports improvements in all three generic searching techniques. How should we move forward on the question of the relationship between the Bradford Distribution and searching techniques?

In teaching undergraduates, I have observed that it is common for them to expect to find resources or websites that provide all the relevant information for their paper in a single location, exactly matching the topic of their paper. In other words, the naive assumption is that chunks of information are perfectly self-contained and complete. At the same time, I have found that beginners often have no idea how to start finding information, and, appear to believe that their one perfect chunk of information may be literally anywhere, that there is no rational way to find it.

Both of these assumptions are inaccurate. Having a generalized understanding of the Bradfordian distribution of all information may provide the basic grounding for a realistic search. As librarians rarely understand this distribution either, we are not always in the best position to help.

What has long been an article of faith among librarians is that all information is findable through the mechanisms of indexing, cataloging, and classification, and that the good searcher invariably utilizes these carefully developed intermediary bibliographic resources. Information literacy instruction for non-librarians is almost always centered around bringing the user into fluency with these sorts of resources.

The experienced humanities scholar constitutes an interesting contrast to the college student and the librarian. The scholar does not have librarian training, but is nonetheless familiar with a number of library research techniques, learned in the course of acquiring the Ph.D. The information seeking research indicates that these scholars rely heavily on browsing and on following citation links (footnote chasing). They rarely, if ever use periodical index databases, and tend to use catalogs only for searches for known items (Stone, 1982; Watson-Boone, 1994; Wiberley, 1989).

The naive user, say, the beginning college student, thus has two models to follow: the librarians, who encourage use of every intermediary access device, from catalogs to online databases, and their professors, who rely most heavily on browsing and footnote chasing.

Another historical metaphor may be suitable at this point. We may view the primary use of browsing and citation linking as an "artisan" approach to searching, that is, a method that dates from the age of skilled craftsmanship. Scholarly research methods have been passed down from mentors to protégés for centuries. By keeping their scholarly focus narrow and deep, humanities researchers have been able to continue that artisan tradition. By searching in a narrow area, the scope of the intellectual territory is kept small enough to allow the continuation of such individualized techniques.

The nineteenth century explosion of new methods of access through classifications, subject indexes, and card catalogs, constituted the beginning of the "industrialization" of searching. With the aid of such devices, one may search vastly more resources in a given period of time (the advantage of "mass production"), but without the personalization made possible by the countless individual search move decisions made in a browsing/footnote-chasing searching technique.

Librarians professionalized in the nineteenth century, at a time when new (Region 2) cataloging and classification techniques were flourishing. Perhaps for that reason, the field created a fused identity with Region 2 devices. Librarians have thus operated on the assumption that all regions of the Bradford distribution of a subject area are best accessed by Region 2 devices, such as catalogs and indexes. In the meantime, scholars and general users tended to muddle along with browsing and footnote chasing as their primary retrieval mechanisms.

Finally, in the late twentieth century, we arrived at the beginning of the Networked Age. When the Web came along, scholars expanded their searching to following Web links, because this technique had a lot of similarities with the footnote chasing they were already familiar with. There have been impressive developments in the potential of networked retrieval to date, but we are still very early in the development of information searching design through networking.

In the end, I believe we should have all three broad types of searching available to us. Work by craftsmen was not abandoned with the advent of the Industrial Revolution, and mass production has not been given up in the Networked Age. We will use them all, with ever more powerful technologies, and ever greater flexibility and effectiveness in our searching. Further, by designing systems to facilitate all three information access methods, effective searching can be supported for all literature sizes and regions.

At the same time, having an understanding of the dynamics of literature distributions may enable us as searchers to make better and more sophisticated decisions about how we want to search and where. It may also prove valuable to test the general conjecture made here, that areas with high numbers of topically relevant materials (relative to all materials in the area) are best searched by browsing, areas with middling numbers of topically relevant items are best searched by directed searching on information-organizational structures, and areas with very sparse ("needle in a haystack") numbers of relevant items are best searched by using links.

REFERENCES

Bates, M.J. (1979). Information search tactics. *Journal of the American Society for Information Science, 30*(4) 205–214.

Bates, M.J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review, 13*(5), 407–424.

Bradford, S.C. (1948). *Documentation*. London: Crosby Lockwood.

Chang, S.-J., & Rice, R.E. (1993). Browsing: A multidimensional framework. In M.E. Williams (Ed.), *Annual review of information science and technology, 28,* 231–276. Medford, NJ: Learned Information.

Chen, Y.S. & Leimkuhler, F.F. (1986). A relationship between Lotka's Law, Bradford's Law, and Zipf's Law. *Journal of the American Society for Information Science, 37*(5), 304–314.

Ellis, D. (1989). A behavioural approach to information retrieval system design. *Journal of Documentation, 45*(3), 171–212.

Fedorowicz, J. (1982). The theoretical foundation of Zipf's Law and its application to the bibliographic database environment. *Journal of the American Society for Information Science, 33*(5), 285–293.

Hood, W.W., & Wilson, C.S. (2001). The scatter of documents over databases in different subject domains: How many databases are needed? *Journal of the American Society for Information Science and Technology, 52*(14), 1242–1254.

Kleinberg, J.M. (1999). Hubs, authorities, and communities. *ACM Computing Surveys, 31* (supp. 4), U21–U23.

Kwasnik, B.H. (1992). *Descriptive study of the functional components of browsing.* Paper presented at the Proceedings of the IFIP TC2\WG2.7 working conference on engineering for human-computer interaction, Elivuoi, Finland, August 10–14, 1992.

Lancaster, F.W., Gondek, V., McCowan, S., & Reese, C. (1991). The relationship between literature scatter and journal accessibility in an academic special library. *Collection Building, 11*(1), 19–22.

Leimkuhler, F.F. (1977). Operational analysis of library systems. *Information Processing & Management, 13*(2), 79–93.

Lempel, R., & Moran, S. (2001). SALSA: The stochastic approach for link-structure analysis. *ACM Transactions on Information Systems, 19*(2), 131–160.

Pontigo-Martinez, J. (1984). Qualitative attributes and the Bradford Distribution. (Unpublished doctoral Dissertation). Champaign-Urbana: University of Illinois.

Price, D.J. de Solla. (1961). *Science since Babylon.* New Haven, CT: Yale University Press.

Price, D.J. de Solla. (1963). *Little science, big science.* New York: Columbia University Press.

Sandstrom, P.E. (1994). An optimal foraging approach to information seeking and use. *The Library Quarterly, 64*(4), 414–449.

Senders, J.W. (1963). Information storage requirements for the contents of the world's libraries. *Science, 141,* 1067–1068.

Stone, S. (1982). Progress in documentation: Humanities scholars: Information needs and uses. *Journal of Documentation, 38*(4), 292–312.

Watson-Boone, R. (1994). The information needs and habits of humanities scholars. *RQ, 34*(2), 203–216.

White, H.D. (1981). 'Bradfordizing' search output: How it would help online users. *Online Review, 5,* 47–54.

Wiberley, S.E. Jr., & Jones, W.G. (1989). Patterns of information seeking in the humanities. *College & Research Libraries, 50*(6), 638–645.

Wilson, A.G., & Bennett, R.J. (1985). *Mathematical methods in human geography and planning.* New York: Wiley.

Wilson, C.S. (1999). Informetrics. *Annual review of information science and technology, 34,* 107–247. Medford, NJ: Information Today.