

## **NGHIÊN CỨU ĐỊNH LƯỢNG DÒNG TIN VỀ LUẬN ÁN THÔNG TIN HỌC, THƯ VIỆN HỌC VÀ THƯ MỤC HỌC DỰA VÀO CƠ SỞ DỮ LIỆU SCIENCE DIRECT (1941-2010)**

**Iu. N. Klimov**

**N**gày nay, không những các nhà nghiên cứu trong nước, danh mục các tác giả và công trình được nêu trong tài liệu tham khảo [2], mà cả các nhà nghiên cứu nước ngoài [5-7] đều chú ý phân tích định tính và định lượng dòng tin khoa học-kỹ thuật. Tuy nhiên, việc nghiên cứu dòng luận án và thông tin về chúng không được tiến hành đầy đủ, ngoài công trình nghiên cứu về thư viện học [4].

Các tài liệu trong nước có giới thiệu một tổng quan mục lục điện tử của các trung tâm thông tin khoa học lớn nhất ở trong và ngoài nước, trong đó, đã đưa ra thông tin về Thư viện quốc gia Nga, Thư viện dân tộc Nga, Thư viện Quốc hội Hoa Kỳ, Thư viện Anh, và nhắc tới nhà xuất bản Elsevier với sản phẩm điện tử Science Direct, chứa hơn 5 triệu bài toàn văn lấy từ 1800 tạp chí. Có thể tìm kiếm thông tin khoa học-kỹ thuật cần thiết theo 25 hướng đề tài và hơn 60 triệu bài tóm tắt [3].

Mục đích của công trình nghiên cứu này là, trên cơ sở phân tích định lượng thông tin dòng bài đăng và thông tin về luận án liên quan đến thông tin học, thư viện học và thư mục học và việc tìm chúng theo từ khóa trong cơ sở dữ liệu Science Direct (1941-2010), xác định vị trí của thông tin học trong lĩnh vực các khoa học thông tin.

Như vậy, tổng số luận án trong trang web Science Direct dao động từ 3 (năm 1941) đến 10.890 (năm 2009), về thông tin học – từ 22 (năm 1950) đến 6.106 (năm

2009), về thư viện học – từ 2 (năm 1950) đến 863 (năm 2009), và về thư mục học – từ 5 (năm 1950) đến 277 (năm 2009).

Trong đó, tổng số thông tin về luận án trong các lĩnh vực đã nêu là: tổng số luận án trên trang web Science Direct - 227.854 (cho năm 2010), về thông tin học - 98.982 (cho năm 2010), về thư viện học - 12.772 (cho năm 2010), và về thư mục học - 7438 (năm 2009).

Như vậy, căn cứ vào tổng số bài đăng về luận án, có thể hình dung hướng đề tài được xem xét, bắt đầu từ số lượng lớn nhất, như sau: tổng số luận án trên trang web Science Direct > thông tin học > thư viện học > thư mục học, nghĩa là thông tin học chiếm vị trí quan trọng giữa thư viện học và thư mục học.

Trong Bảng 1, có thể thấy sự phân bố theo thời gian của các xuất bản phẩm định kỳ và tiếp tục, chứa các thông tin về luận án theo các hướng đề tài được xem xét trên trang web Science Direct, nơi cho thấy tổng số thông tin về luận án, so với các thời kỳ nghiên cứu tất cả các loại hình xuất bản phẩm trước đây, đang tăng lên ở tạp chí, sách và bài tạp chí. Do đó, tính theo số lượng tạp chí, mức tăng này là 11,8 lần so với thời kỳ 1941-1960, 1991-2010 (tháng 3), và tăng 4,7 lần so với giai đoạn 1981-1990. Trong những thời kỳ này, số lượng sách thay đổi từ 10,6 đến 7 lần, còn bài tạp chí từ 11,8 lần đến 5,4 lần.

## *Nhìn ra thế giới*

**Bảng 1:** Sự phân bố theo thời gian của các xuất bản phẩm định kỳ và tiếp tục, chứa các thông tin về luận án thông tin học, thư viện học và thư mục học, trên trang web Science Direct

Hướng đề tài	Thời kỳ	Tạp chí	Sách	Bài
Tổng số thông tin về luận án	1941 - 1960	3639	154	3641
	1961 - 1980	42877	1636	42901
	1981 - 1990	49048	1649	42910
	1991 - 2010 (Tháng ba)	228180	11627	233859
Thông tin học	1950 - 1960	443	55	443
	1961 - 1970	2889	312	2893
	1971 - 1990	27376	1426	27441
	1991 - 2010 (Tháng ba)	96218	6850	99428
Thư viện học	1949 - 1971	378	34	378
	1969 - 1991	3003	184	3010
	1991 - 2010 (Tháng ba)	12353	1246	13055
	Thư mục học	1949 - 1971	601	104
1969 - 1991		3011	251	3025
1991 - 2010 (Tháng ba)		7147	1188	7803

Nếu xem xét tỷ lệ các loại hình xuất bản phẩm có chứa thông tin về luận án *thông tin học*, trong thời kỳ 1991-2010 (tháng ba) so với 1950-1960, thì mức tăng của tạp chí là 6,5 lần, còn trong thời kỳ 1991-2010 (tháng ba) so với 1971-1990, là 3,5 lần. Trong những thời kỳ này, số lượng sách thay đổi từ 5,7 lần đến 4,8 lần, còn bài tạp chí từ 6,5 lần đến 3,6 lần.

Tỷ lệ các loại hình xuất bản phẩm có chứa thông tin về luận án *thư viện học* trong thời kỳ 1991-2010 (tháng ba) so với 1949-1971, tăng 7,9 lần, còn trong thời kỳ 1991-2010 (tháng ba) so với 1969 tăng 4,1 lần. Số lượng sách trong những thời kỳ này thay đổi từ 5,1 lần đến 6,8 lần và đối với bài tạp chí, từ 8 lần đến 4,3 lần.

Tỷ lệ các loại hình xuất bản phẩm chứa thông tin về luận án *thư mục học*

## *Nhìn ra thế giới*

cho thấy, trong thời kỳ 1991-2010 (tháng ba) so với 1949-1971, tăng 5 lần, còn so với thời kỳ 1969-1991, tăng 2,4 lần. Số lượng sách trong những thời kỳ này thay đổi từ 2,4 lần đến 4,7 lần và đối với bài tạp chí, từ 5 lần đến 4,7 lần.

Tuy nhiên, nếu xem xét tỷ lệ số lượng xuất bản phẩm được nghiên cứu từ những dữ liệu ban đầu trong từng trường hợp riêng, thì giá trị của chúng còn cao hơn nhiều.

Ngoài ra, người ta cũng xác định được tỷ lệ số lượng các bài đăng về luận án theo thập niên: 1960-1969, 1970-1979, 1980-1989, 1990-1999 và 2000-2009 đã giảm đi so với tổng số thông tin về luận án từ 4,4 lần xuống còn 1,4 lần, đối với thông tin học – từ 6,6 lần xuống còn 1,5 lần, đối với thư viện học - từ 2,5 lần xuống còn 1,7 lần; và đối với thư mục học – từ 2 lần xuống còn 1,2 lần, mặc dầu số lượng bài đăng vẫn tăng theo thập niên.

**Bảng 2a:** *Mô hình hóa số lượng tích hợp các bài đăng theo tổng số thông tin về luận án thông tin học, thư viện học và thư mục học và theo các luận án về thư viện học*

Đề tài	Quan hệ phụ thuộc					
	Lũy thừa $y=ax^b$			Thuật toán $y = alnx + b$		
	$a$	$b$	$R^2$	$a$	$b$	$R^2$
Thông tin về luận án nói chung trên trang web Science Direct, 1955-2009	79,79	1,85	0,94	58854,00	- 10152,00	0,57
Thông tin về luận án thông tin học trên trang web Science Direct, 1950-2010	2,70	2,46	0,95	22727,00	-46372,00	0,49
Thông tin về luận án thư viện học trên trang web Science Direct, 1950-2010	0,77	2,23	0,96	2732,00	-5604,90	0,46
Thông tin về luận án thư mục học trên trang web Science Direct, 1950-2009	3,28	1,84	0,98	2010,10	-3831,90	0,60
Các luận án 1903-2004 theo Romanov P.X., 2006 [4]	0,14	2,12	0,90	695,93	1228,60	0,43
Thông tin về luận án thư viện học trên trang web Science Direct, 1950-2000	1,02	2,11	0,96	1554,70	-2946,30	0,46
Các luận án 1950-2000 theo Romanov P.X., 2006 [4]	16,61	1,20	0,86	769,19	-1289,10	0,60

## *Nhìn ra thế giới*

**Bảng 2b:** Mô hình hóa số lượng tích hợp các bài đăng theo tổng số thông tin về luận án thông tin học, thư viện học và thư mục học và theo các luận án thư viện học

Đề tài	Quan hệ phụ thuộc						
	Hàm số mũ $y=a.e^{bx}$			Đa thức bậc 2 $y = ax^2 + bx+c$			
	$a$	$b$	$R^2$	$a$	$b$	$c$	$R^2$
Thông tin về luận án nói chung trên trang web Science Direct, 1955-2009	1455,70	0,096	0,94	76,60	-1001,70	4858,00	0,99
Thông tin về luận án thông tin học trên trang web Science Direct, 1950-2010	140,30	0,12	0,93	41,02	-1034,70	5377,30	0,99
Thông tin về luận án thư viện học trên trang web Science Direct, 1950-2010	27,15	0,11	0,94	5,57	-161,44	955,98	0,99
Thông tin về luận án thư mục học trên trang web Science Direct, 1950-2009	66,32	0,09	0,90	2,22	-6,87	37,60	0,99
Các luận án 1903-2004 theo Romanov P.X., 2006 [4]	6,19	0,08	0,97	0,84	-34,93	272,87	0,99
Thông tin về luận án thư viện học trên trang web Science Direct, 1950-2000	21,35	0,13	0,94	4,53	-112,82	590,42	0,99
Các luận án 1950-2000 theo Romanov P.X., 2006 [4]	80,88	0,08	0,98	1,22	-6,57	88,14	0,99

Để nghiên cứu thuộc tính số lượng tích hợp các bài đăng theo tổng số thông tin về luận án, luận án thông tin học, thư viện học và thư mục học trong thời kỳ 1950-2010 (tổng số thông tin về luận án) và 1941-2009 (các bài đăng về luận án thông tin học, thư viện học và thư mục học), người ta đã tiến hành mô hình hóa số lượng bài đăng theo quan hệ phụ thuộc lũy thừa,

thuật toán, hàm số mũ và đa thức bậc 2 (Bảng 2 a-b).

Việc áp dụng *quan hệ phụ thuộc lũy thừa* cho thấy, theo trị số  $a$  (số lượng bài đăng ban đầu), thông tin về luận án phân bố như sau: tổng số bài đăng về luận án, về luận án thư mục học, thông tin học và thư viện học 79,79; 3,28; 2,70 và 0,77.

Tốc độ tương quan của số lượng tích

## Nhìn ra thế giới

hợp các bài đăng (trị số  $b$ ) là lớn nhất đối với các bài đăng về luận án thông tin học (2,46) và nhỏ nhất – đối với các bài đăng thư mục học (1,84), với hệ số tính toán tương quan  $R^2$  từ 0,98 (thư mục học) đến 0,96 (thư viện học).

Việc mô hình hóa các dòng tích hợp được nghiên cứu theo *quan hệ phụ thuộc thuật toán* cho thấy, theo trị số lớn nhất  $a$  (số lượng bài đăng ban đầu) thông tin về các luận án được phân bố như sau: tổng số bài đăng về luận án, về luận án thông tin học, thư viện học và thư mục học là 58854.00, 22727.00, 2732.00, và 2010.00. Trong trường hợp này, mức gia tăng số lượng tích hợp các bài đăng (trị số  $b$ ) trong mọi trường hợp đều âm. Hệ số lớn nhất đa tương quan  $R^2$  bằng 0,60 (thông tin về luận án thư mục học), còn hệ số nhỏ nhất là 0,46 (thông tin về luận án thư viện học).

Mức gia tăng số lượng tích hợp các bài đăng là âm đối với tổng số thông tin về luận án, đối với thông tin về luận án thông tin học, thư viện học và thư mục học: - 110152,00, - 46372,00, - 5604,90 và - 3831,90 các bài đăng. Các hệ số đa tương quan được tính toán đều thấp.

Việc áp dụng *quan hệ phụ thuộc hàm số mũ* để mô hình hóa số lượng tích hợp bài đăng cho thấy, số lượng bài đăng một năm (trị số  $a$ ) dao động từ 1455,70 (tổng số thông tin về luận án) đến 27,15 (thông tin về luận án thư viện học) với tốc độ hàm số mũ tương quan của số lượng tích hợp bài đăng từ 0,12 (thông tin về luận án thông tin học) đến 0,09 (tổng số thông tin về luận án và thông tin về luận án thư mục học) với hệ

số đa tương quan  $R^2$  từ 0,94 (tổng số thông tin về luận án và bài đăng về luận án thư viện học) đến 0,90 (thông tin về luận án thư mục học).

So sánh tốc độ tương quan theo *quan hệ phụ thuộc hàm số mũ của số lượng tích hợp bài đăng* với tốc độ tương quan theo *quan hệ phụ thuộc lũy thừa* đã phát hiện được xu thế chung là: trị số nhỏ nhất của chúng tương ứng với các thông tin về luận án nói chung và thông tin về luận án thư mục học, còn trị số lớn nhất tương ứng với các thông tin về luận án thông tin học và thư mục học đối với khối lượng khá lớn các dữ liệu và các thời kỳ lớn công bố chúng trong các xuất bản phẩm định kỳ và tiếp tục.

Mô hình hóa số lượng tích hợp bài đăng về luận án *theo đa thức bậc 2* cho thấy, theo trị số  $a$  với  $x^2$  các bài đăng về luận án phân bố như nhau, cũng như theo *quan hệ phụ thuộc thuật toán*, nghĩa là trị số lớn nhất  $a$  tương ứng với tổng số bài đăng về luận án thư mục học (78,60 và 2,22 bài đăng). Trị số  $b$  với  $x$  trong mọi trường hợp đều là âm, nhưng xu hướng đã nêu đối với các phần phân bố trong số lượng tích hợp các bài đăng vẫn như trước.

Đối với vết tự do của phương trình  $c$ , xu hướng chung nêu trên vẫn được bảo toàn. Đồng thời đối với số lượng tích hợp các bài đăng về luận án thư mục học, xu hướng ấy là âm. Hệ số tính toán đa tương quan  $R^2$  trong mọi trường hợp nêu trên đều bằng 0,99.

### **Động thái và số lượng tích hợp của các luận án thư viện học bằng tiếng Anh**

Việc nghiên cứu động thái số lượng các

## Nhìn ra thế giới

luan án thư viện học bằng tiếng Anh trong thời kỳ 1903-2004 dựa trên công trình nghiên cứu của P.X.Pômanov [4] rất đáng chú ý, ở đó, chúng tôi thấy số lượng luận án tăng không đều trong những năm 1958, 1960, 1963, 1966, 1970, 1972, 1975, 1977, 1983, 1987, 1989, 1991-1992 và 1995-1996. Tổng số luận án là 3130. 50% tổng số luận án là vào năm 1987.

Theo *quan hệ phụ thuộc lũy thừa*, số lượng tích hợp luận án trong năm (trị số  $a$ ) là 695,93, còn mức độ gia tăng là 1228,60 luận án với hệ số đa tương quan  $R^2 = 0,43$ .

Mô hình hóa số lượng tích hợp luận án theo *quan hệ phụ thuộc hàm số mũ* cho thấy, số lượng luận án trong năm bằng 6,19 với tốc độ hàm số mũ tương quan là 0,08 với hệ số tính toán đa tương quan  $R^2 = 0,97$ .

Việc áp dụng *đa thức bậc 2* phát hiện ra rằng, trị số  $a$  với  $x^2$  bằng 0,84 luận án, trị số  $b$  với  $x$  là âm (-34,93), vết tự do của phương trình  $c$  bằng 272,87 luận án với hệ số tính toán đa tương quan  $R^2 = 0,99$ .

### **Phân tích so sánh mô hình hóa số lượng tích hợp các bài đăng luận án dựa vào trang web Science Direct và số lượng tích hợp luận án thư viện học trong thời kỳ 1950-2000**

Sẽ là hợp lý khi chỉ ra sự giống nhau và khác nhau của các chỉ tiêu định lượng số lượng tích hợp các bài đăng luận án trên trang web Science Direct và số lượng tích hợp các bài đăng về thư viện học nói riêng.

So sánh mô hình hóa theo *quan hệ phụ thuộc lũy thừa* số lượng tích lũy các bài đăng về luận án thư viện học dựa vào trang

web Science Direct và số lượng tích hợp các luận án trong thời kỳ 1950-2000 cho thấy, theo trị số  $a$ , số lượng ban đầu của các bài đăng và luận án khác hẳn nhau.

Do đó, theo trị số  $a$ , số lượng các bài đăng một năm về luận án và bản thân các luận án nói riêng đã thấy có sự khác biệt sau đây: đối với các bài đăng về luận án, sự khác biệt là 1,01 luận án, còn đối với các luận án thì là 16,61 với tốc độ tương quan trị số  $b$  của các thông tin về luận án là 2,11, còn đối với các luận án là 1,20, hệ số đa tương quan  $R^2 = 0,96$  và 0,86.

Mô hình hóa theo *quan hệ phụ thuộc thuật toán* số lượng tích hợp các bài đăng về luận án thư viện học dựa vào trang web Science Direct và số lượng tích hợp các luận án trong thời kỳ 1950-2000 theo trị số  $a$  của số lượng các bài đăng trong một năm về luận án và bản thân các luận án nói riêng đã phát hiện những sự khác biệt sau đây: đối với các bài đăng về luận án, trị số đó là 1554,70, đối với các luận án là 769,19, trị số  $b$  đối với cả hai trường hợp là âm, còn các hệ số đa tương quan  $R^2 = 0,46$  và 0,60.

Mô hình hóa theo *quan hệ phụ thuộc hàm số mũ* số lượng tích hợp các bài đăng về luận án thư viện học dựa vào trang web Science Direct và số lượng tích hợp các luận án trong thời kỳ 1950-2000 theo trị số  $a$  của số lượng các bài đăng trong một năm về luận án và bản thân các luận án nói riêng đã phát hiện những sự khác biệt sau đây: đối với các bài đăng về luận án, trị số đó là 21,35, đối với các luận án là 80,88, còn tốc độ hàm số mũ tương quan là 0,13

và 0,08, với các hệ số đa tương quan  $R^2 = 0,94$  và  $0,98$ .

Việc áp dụng đa thức bậc 2 để nghiên cứu số lượng tích hợp các bài đăng về luận án thư viện học dựa vào trang web Science Direct và số lượng tích hợp các luận án trong thời kỳ 1950-2000 cho thấy, trị số  $a$  với  $x^2$  bằng 4,53 đối với các bài đăng về luận án, đối với các luận án là 1,22, trị số  $b$  với  $x$  trong cả hai trường hợp đều là âm (-6,57 các bài đăng về luận án và -112,82 các luận án), vế tự do của phương trình  $c$  tương ứng bằng 88,14 và 590,42 với hệ số tính toán đa tương quan  $R^2$  mỗi loại là 0,99.

Như vậy, việc phân tích so sánh mô hình hóa các dòng bài đăng về luận án và bản thân các luận án nói riêng [4] trong cùng một thời kỳ cho thấy sự giống nhau và khác nhau giữa chúng: cả hai dòng được nghiên cứu đều xấp xỉ gần bằng nhau theo đa thức bậc 2, các quan hệ phụ thuộc hàm số mũ và lũy thừa, chỉ khác nhau ở số lượng tích hợp các bài đăng về luận án và bản thân các luận án nói riêng với tốc độ tương quan khác nhau; tốc độ này cao hơn 1,76 lần đối với các bài đăng về luận án trên trang web Science Direct, so với bản thân các luận án.

Sự khác biệt về tốc độ hàm số mũ tương quan của số lượng tích hợp các bài đăng về luận án trên trang web Science Direct và bản thân các luận án cao hơn 1,62 lần đối với số lượng tích hợp các bài đăng về luận án so với bản thân các luận án.

**Một vài thuộc tính của các bài đăng liên quan đến thông tin học, thư viện học và thư mục học**

Căn cứ vào những dữ liệu thu được trên trang web Science Direct, hiện nay có sự tăng chậm lại số lượng các bài đăng theo thập niên. Chẳng hạn, nếu trong thời kỳ 1960-1969, tổng số thông tin về luận án so với thời kỳ 1950-1959 tăng gấp 4,4 lần, thì trong thời kỳ 2000-2009 so với thời kỳ 1990-1999 chỉ tăng 1,4 lần, còn các bài đăng về luận án thông tin học, thư viện học và thư mục học thì mức tăng tương ứng là 6,6 và 1,5; 2,5 và 1,7; 2,0 và 1,2 lần.

Số lượng các luận án thư viện học trong thời kỳ 1960-1969 so với thời kỳ trước 1950-1959 tăng 3 lần; còn trong thời kỳ 1990-2000 so với 1980-1990 chỉ tăng 1,4 lần.

Ngoài ra, cần chú ý rằng, các dòng tích hợp được nghiên cứu khẳng định sự có mặt của ba khu vực không đồng đều đối với các bài đăng các thông tin chung về luận án, đối với các thông tin về luận án thông tin học, thư viện học và thư mục học, và cả đối với bản thân các luận án thư viện học dựa vào các dữ liệu của công trình nghiên cứu [4].

**Các xuất bản phẩm định kỳ và tiếp tục chủ yếu về thông tin học, thư viện học và thư mục học dựa vào trang web Science Direct**

Nghiên cứu các xuất bản phẩm định kỳ và tiếp tục chủ yếu, trong đó có chứa thông tin về luận án thông tin học, thư viện học và thư mục học trên trang web Science Direct cho thấy, toàn bộ dòng tài liệu thông tin theo tổng số thông tin về luận án có chứa 28.965 bài đăng từ 228.180 tạp chí, theo các bài đăng thông tin về luận án

## Nhìn ra thế giới

thông tin học là 12.034 bài đăng từ 96.218 tạp chí, và về thư viện học là 2.448 bài đăng từ 12.353 tạp chí.

Những xuất bản phẩm định kỳ và ra tiếp tục hàng đầu tính theo tổng số bài đăng về luận án (hơn 100 bài) có 33 tạp chí (44.403 bài, 19,0% tổng số bài), về thông tin học có 20 tạp chí (12.034 bài, 12,1% tổng số bài), về thư viện học có 9 tạp chí (1631 bài, 12,5% tổng số bài) và về thư mục học có 6 tạp chí (929 bài, 11,9% tổng số bài).

### Đề xuất khu vực hạt nhân của các tạp chí và định luật Bradford

Rõ ràng là đối với từng lĩnh vực chủ đề, do các người dùng tin chủ yếu đưa ra, có thể xác định số lượng tối đa các tạp chí hoặc sách theo định luật Bradford [9].

Việc xác định khu vực hạt nhân của các tạp chí theo các thông tin chung về luận án trên trang web Science Direct được tiến hành theo hai quan hệ phụ thuộc đại số: *quan hệ hàm số mũ và đa thức bậc 2*. Đối với tổng số thông tin về luận án (TTLA), *quan hệ phụ thuộc hàm số mũ* có dạng sau đây:  $y_{TTLA} = 14829,00x \cdot e^{0,03x}$ ,  $R^2=0,60$  và  $y_{TTLA} = - 31,39 x^2 + 2311,40x + 3684,20$ ,  $R^2=0,99$ . Mỗi quan hệ này được thể hiện bởi họ hàng các xuất bản phẩm định kỳ và tiếp tục: Journal of Personality and Social Psychology, Geochimica et Cosmochimica Acta, Vật liệu Polyme (Liên Xô/Nga), Lancet, Psychological Bulletin, Social Science & Medicine và European Journal of Operational Research.

Khác với công trình nghiên cứu [9], trong đó nêu ra ba khu vực đồng đều phân bố xuất bản phẩm, ở đây phát hiện sự có

mặt của ba khu vực không đồng đều: khu vực hạt nhân thứ nhất chứa 7 tạp chí và 12.541 bài, khu vực thứ hai chứa 29 tạp chí và 22096 bài, khu vực thứ ba chứa 10 tạp chí và 695 bài với tỷ lệ tương quan 1:2, 04:1,02.

Khu vực hạt nhân được đề xuất trên cơ sở mối quan hệ phụ thuộc hàm số mũ và đa thức bậc 2:  $y_{TTLA} = 2370,50 \cdot e^{0,09x}$ ,  $R^2=0,81$  và  $y_{TTLA} = - 14,20x^2 + 844,21x + 599,88$ ,  $R^2=0,99$  xuất bản phẩm định kỳ và tiếp tục theo các thông tin về luận án thông tin học, khu vực này được đại diện bởi hai xuất bản phẩm định kỳ và tiếp tục: Journal of Personality and Social Psychology và Social Science & Medicine, mà quy luật Bradford về sự có mặt của ba khu vực khẳng định điều này. Chúng tôi đã chỉ rõ sự phân bố không đồng đều số lượng tạp chí và các bài đăng trong các khu vực này: khu vực thứ nhất là khu vực hạt nhân chứa 2 tạp chí và 2104 bài, khu vực thứ hai chứa 13 tạp chí và 7803 bài, khu vực thứ ba chứa 5 tạp chí và 2127 bài với tỷ lệ tương quan 1:4, 70:1,22.

Trên cơ sở quan hệ phụ thuộc hàm số mũ và đa thức bậc 2:  $y_{TTLATV} = 613,37 \cdot e^{0,08x}$ ,  $R^2=0,75$  và  $y_{TTLATV} = -4,36 x^2 + 195,13x + 209,84$ ,  $R^2=0,99$ , ba tạp chí: Information, Processing & Management, Library & Information Science Research và Lancet tạo thành khu vực hạt nhân các xuất bản phẩm định kỳ và tiếp tục theo các thông tin về luận án thư viện học (TTLATV), mà quy luật Bradford về sự có mặt của ba khu vực khẳng định điều này. Chúng tôi đã chỉ rõ sự phân bố không đồng đều số lượng tạp

## Nhìn ra thế giới

chí và các bài đăng trong các khu vực này: khu vực hạt nhân thứ nhất chứa 3 tạp chí và 528 bài, khu vực thứ hai chứa 12 tạp chí và 1646 bài, khu vực thứ ba chứa 5 tạp chí và 271 bài với tỷ lệ tương quan 1:4, 12:1,12.

Trên cơ sở *quan hệ phụ thuộc hàm số mũ và đa thức bậc 2*:  $y_{TTLATM} = 404,77 \cdot e^{0,09x}$ ,  $R^2=0,77$  và  $y_{TTLATM} = -3,20x^2 + 151,66x + 13x + 82,98$ ,  $R^2=0,99$ , đối với các thông tin về luận án thư mục học (TTLATM), ba tạp chí: Psychological Bulletin, Accounting, Organization and Society và Lingua tạo thành khu vực hạt nhân các xuất bản phẩm định kỳ và tiếp tục, mà quy luật Bradford về sự có mặt của ba khu vực khẳng định điều này. Chúng tôi đã chỉ rõ sự phân bố không đồng đều số lượng tạp chí và các bài đăng trong các khu vực này: khu vực hạt nhân thứ nhất chứa 3 tạp chí và 506 bài, khu vực thứ hai chứa 12 tạp chí và 1907 bài, khu vực thứ ba chứa 5 tạp chí và 281 bài với tỷ lệ tương quan 1:3, 17:1,18.

Việc áp dụng quy luật Bradford đã được xem xét để phát hiện các từ được sử dụng nhiều nhất khi tìm kiếm các thông tin về luận án trên trang web Science Direct.

Việc nghiên cứu sự phân bố tích hợp các từ khóa mô tả các dòng thông tin chủ đề về luận án thông tin học, thư viện và thư mục học cho thấy, số lượng tích hợp các từ khóa về thông tin theo tổng số luận án trên trang web Science Direct là 47 với tần số 18081. Việc xác định các từ khóa chính được thực hiện theo hai quan hệ phụ thuộc đại số: hàm số mũ và đa thức bậc 2.

Số lượng tích hợp từ khóa cho thông tin về luận án (TTLA) trên trang web Science

Direct tính theo *quan hệ phụ thuộc hàm số mũ và đa thức bậc 2* có dạng sau đây:  $y_{TTLA} = 6237,50 \cdot e^{0,03x}$ ,  $R^2=0,60$  và  $y_{TTLA} = -13,06x^2 + 928,22x + 1996,50$ ,  $R^2=0,98$ .

Bảy từ khóa đã được phát hiện (nghĩa là khu vực thứ nhất): tâm lý học xã hội, trẻ con, giao diện, tâm thần, phần mềm, tiếng Anh và định lý với tần số 7995. 32 từ khóa với tần số tích hợp 9773 tạo thành khu vực thứ hai, 12 từ khóa với tần số tích hợp 313 tạo thành khu vực thứ 3 với tỷ lệ tương quan 1:2, 21:1,02; thông tin về luận án thông tin học theo quan hệ phụ thuộc hàm số mũ và đa thức bậc 2 có dạng sau đây:  $y_{TTLATT} = 1023,00 \cdot e^{0,10x}$ ,  $R^2=0,83$  và  $y_{TTLATT} = -5,71x^2 + 370,46x + 233,41$ ,  $R^2=0,99$ . Ba từ khóa được sử dụng nhiều nhất. (khu vực 1) là tâm lý học xã hội, trẻ em và giao diện với tần số 1311; 12 từ khóa với tần số 3160 tạo thành khu vực thứ 2; 5 từ khóa với tần số 942 tạo thành khu vực thứ 3 với tỷ lệ tương quan 1:3, 41:1,21.

Số lượng tích hợp các từ khóa cho thông tin về luận án thư viện học trên trang web Science Direct tính theo *quan hệ phụ thuộc hàm số mũ và đa thức bậc 2* có dạng sau đây:  $y_{TTLATV} = 220,44 \cdot e^{0,09x}$ ,  $R^2=0,86$  và  $y_{TTLATV} = -0,01x^2 + 70,64x + 68,74,50$ ,  $R^2=0,99$ . Các từ khóa được sử dụng nhiều nhất. (khu vực 1) là phần mềm, thông tin học và thư viện với tần số 276, 12 từ khóa với tần số 620 tạo thành khu vực thứ hai, 5 từ khóa với tần số 187 tạo thành khu vực thứ 3 với tỷ lệ tương quan 1:3, 25:1,21; thông tin về luận án thư mục học tính theo *quan hệ phụ thuộc hàm số mũ và đa thức bậc 2* có dạng sau đây:  $y_{TTLATM} = 123,54 \cdot e^{0,10x}$ ,  $R^2=0,87$  và  $y_{TTLATM} = -0,48x^2 + 13,41x$

## Nhìn ra thế giới

+ 30,65,  $R^2=0,99$ . Các từ khóa (khu vực 1) là toán học và phần tử cuối với tần số 116, khu vực 2 có 13 từ khóa với tần số 458, khu vực 3 có 5 từ khóa với tần số 135 với tỷ lệ tương quan 1:4, 95:1,24.

Nghiên cứu sự phân bố tích hợp số lượng bài và từ khóa theo tần số nhỏ dần cho thấy, trong mọi trường hợp áp dụng quan hệ phụ thuộc hàm số mũ và đa thức bậc 2 trước điểm bắt đầu và sau điểm kết thúc giao nhau với các đường cong phản ánh các mối quan hệ phụ thuộc này, người ta tách ra ba khu vực tương ứng với vùng hạt nhân và hai vùng phân tán của các bài đăng hoặc từ khóa. Quy luật Bradford đối với các bài đăng và quy luật Zipf đối với việc phân bố tích hợp các từ về sự có mặt của khu vực hạt nhân và hai khu vực phân tán số lượng tích hợp các bài đăng và các từ [8] khẳng định các kết quả đã được trình bày.

Như vậy, các kết quả nghiên cứu định lượng thông tin đối với dòng bài đăng về

luận án thông tin học, thư viện và thư mục học dựa vào cơ sở dữ liệu Science Direct (1941-2010) đã cho phép: 1) Nghiên cứu định lượng động thái và số lượng tích hợp thông tin về tất cả các luận án thông tin học, thư viện và thư mục học, 2) Nghiên cứu động thái và số lượng tích hợp các luận án thư viện học, 3) Phát hiện các xuất bản phẩm định kỳ và tiếp tục quan trọng nhất, 4) tiến hành mô hình hóa số lượng tích hợp của chúng, 5) Áp dụng *quan hệ phụ thuộc hàm số mũ và đa thức bậc 2* để phát hiện chính xác ba khu vực không đồng đều, tương ứng với vùng hạt nhân và hai vùng tán xạ các bài đăng hoặc từ khóa; 6) Chỉ ra dòng thông tin về thông tin học chiếm vị trí nổi bật trong lĩnh vực các khoa học thông tin.

**Vũ Văn Sơn (Dịch)**

**Nguồn:** *Nauchno-tekhnicheskaja informaxija. Serija 1, Organizaxija i metodeka informaxionnoj raboty, 2010, No. 12, Page. 27-32*

### Tài liệu tham khảo

1. Горькова В.Н. Информетрия // Итоги науки и техники. - М.: ВИНТИ, 1998 . – Т 10 . – 190с.
2. Климов Ю.Н. Научная продуктивность и публикационная активность ученых-физиков в базе данных SCOPUS //ИТИ. Сер.1. - 2010. – N 4. – С.12-18.
3. Фицер А.М. Обзор электронных каталогов крупнейших зарубежных информационных центров //ИТИ. Сер.1. - 2004. – N 9. – С.26-34.
4. Романов П.С. Библиографический список англоязычных диссертаций по библиотековедению (1903-2004гг.) . - М., 2006 . – 376с. (Кабинет библиографии, электронная версия в РГБ) .
5. Efthimidias R. The Growth on the OPAC Literature // Journal of the American Society for Information Science. – 1990.- Vol.41, N.7. – p.342-347.
6. Egghe L., Rao I.K.R. Classification of growth models // Scientometrics. – 1992. – Vol.25. – p.5-46.
7. Egghe L., Rao I.K.R. Citation Age Data and the Obsolescence Function: Fits and Explanations // Information processing and Management. – 1992.- Vol.28, N.2. – p. 201-217.
8. Климов Ю.Н. Моделирование распределения Бредфорда и Ципфа // Межотраслевая информационная служба. - 2009. – N 4(145) . – С.60-62.
9. Солто Дж. Динамические библиотечно-информационные системы . - М.: Мир, 1979 . – С.198.