

KTALIS 도서자료 검색시스템의 이용기능 분포도 및 응답시간을 중심으로한 시스템성능에 대한 고찰

A Study on the Performance of KTALIS Library IR System
by the Usage Pattern and the Response Time

김재일* 김영현* 이해옥* 허문행*

초 록

본고에서는 KTA에서 도서자료 토털시스템의 일부로 개발하여 이용중인 단행본 검색시스템에 대한 사용자의 이용기능 분포도를 축적된 통계자료를 이용하여 분석하고 각 이용기능별 응답시간과 각 이용기능별 검색성공률 및 전체 성공율을 통계적인 도표와 그래프로 작성하여 시스템의 성능을 수치적으로 계량화하여 보았다.

ABSTRACT

In this paper, we analyzed the statistical data logged by the library information retrieval system which was developed in KTA as a part of total system. System performance was measured numerically by means of tables and graphs representing the usage patterns, the response times, and the success rates of each searching item.

1. 서 론

한국전기통신공사에서는 효율적인 정보자료 제공을 위하여 1986년부터 도서정보시스템(K-TALIS)의 개발에 착수하여 1987.7월 1차로 단행본시스템을 개발완료하고 개통하여 도서자료의 처리와 검색기능을 온라인 서비스로 제공

하게 되었다. 본 시스템은 완벽한 한글 온라인 검색시스템으로 사용자에게 최고 수준의 편의를 제공하기 위하여 다양한 내부처리 기능 및 검색 명령어를 구비하여 독자적인 DBMS 형식으로 개발된 것이다.

본고에서는 이러한 시스템에 대한 진단의 일환으로 87.7월부터 약 1년간에 축적 보존된 사

* 한국전기통신공사 사업지원단

용자의 시스템 이용내역을 여러가지 형태로 분석하여 시스템기능 이용성향을 파악하고 검색 명령에 대한 응답시간을 측정함으로써 시스템 자체의 성능을 계량화하여 보았으며 유사 시스템 개발시 참고자료로 활용될 수 있도록 하였다.

2. KTALIS개요

2.1 소프트웨어 시스템의 구성

KTALIS는 크게 4개의 subsystem(단행본, 정기간행물, 출장보고서, 해외망접속)으로 모듈화하여 연차적으로 개발작업이 진행중에 있으며 이중 단행본시스템은 사서의 업무처리를 위한 수서, 목록, 대출반납, 동의어 처리 및 각종 통계관리, 그리고 사용자를 위한 검색 등의 기능을 수행하도록 개발되었으며 독립된 복수의 자료관리부서에서 각기 별도의 자료 DB를 구축하여 기관의 특성을 살릴수 있도록 하였다.

단행본 검색기능에는 9개의 검색대상 항목이 서명, 저자명, 등록번호, 분류기호, 키워드, 언어, 출판년도, 주제어, 내용형식 있으며 이들을 논리연산자를 사용하여 자유로 조합검색을 할 수 있으며 단일 항목에 대해서도 여러가지 검색 key를 조합하여 사용할 수 있고 어떤 범위에 대한 검색도 허용하고 있다. 출판년도와 등록번호는 관계연산자를 사용하여 검색할 수 있다. 검색의 결과는 해당 레코드가 들어있는 위치들로 이루어진 set로 만들어지며 이 set에 대한 검색도 허용하고 있으며 set끼리 연산자를 이용하여 조합할 수도 있다. 검색대상 항목 중 언어와 내용형식은 코드화되어 있으며 등록번호는 일련번호, 분류기호는 LC분류기호이다.

2.2 하드웨어 환경

시스템의 개발 및 운용환경은 AT&T사에서 UNIX System V의 주목표시스템으로 발표한 3B20S Mini-computer이며 Memory 6MB, Disk 300MB 3대, 64개의 터미널 port로 본사 18, 연구부서 24, 전국 각 지방 21개의 분포로 터미널망이 구성되어 있다.

현재 KTALIS에는 136,800 block(68.4MB)의 disk가 할당되어 있는데 약 10만권의 단행본 자료처리가 가능하며 시스템을 KTALIS 전용으로 사용하며 현재의 응답시간 능력을 유지하고자 한다면 대략 50~60만권의 자료처리가 가능할 것이다.

2.3 기 타

현재 터미널을 사용하는 사용자군은 크게 자료관리 업무를 수행하는 4개의 자료실(본사, 사업지원단, 연수원, 품질보증단)군과 연구개발 업무를 수행하는 연구부서군, 그리고 일반 현업 사용자군으로 3분 할 수 있으며 약 3만여권의 단행본 DB가 구축되어 있고 이 자료를 검색한 사용자의 이용현황이 14,797개의 레코드로 축적되어 있다.

3. 검색시스템의 이용 경향

사용자의 검색시스템 이용 경향을 파악하기 위하여 지금까지 수집된 14,797개의 레코드에 들어있는 검색명령을 12가지 유형으로 - 9가지 검색항목을 하나씩만 사용한 9가지 경우, 둘 이상의 검색항목을 동시에 사용한 경우, 검색결과인 set와 검색명령을 조합한 경우, set끼리

조합한 경우 분류하였다. 이 분류에 따라서 검색 성공율, 실패의 원인등을 분석하였으며 검색명령에서 자주 사용된 단어들을 별도로 수집하였다. 검색 결과의 건수가 1 이상일때를 성공으로 취급하였는데 이는 반드시 사용자가 원하는 자료를 찾았음을 의미하는 것은 아니다.

3.1 검색명령 유형별 이용도와 성공율

〈그림 1〉은 검색명령 유형별 이용도를 그림으로 나타낸 것이며 〈표 1〉은 터미날의 설치위치에 따른 검색명령 유형별 이용도와 성공율을 보여 주고 있다. 먼저 검색 성공율에 대하여 살펴보면 전체적으로 61.4%가 성공, 38.6%가 실패로 나타났으며 출판년도로 검색한 경우가 성공율 87.9%로 가장 높았고 주제어로 검색한 경우가 성공율 38.1%로 가장 낮았다.

터미날의 설치위치에 따라서는 연구부서에서 전체적인 성공율 66.9%로 가장 높았고 자료실, 연구부서, 일반부서에서 모두 출판년도로 검색한 경우의 성공율이 가장 높고 주제어로 검색한 경우의 성공율이 가장 낮았다. 출판년도는 구조적으로 성공율이 높을수 밖에 없으며 주제어는 이용자에게 잘알려지지 않은 것으로 분석된다.

사용한 회수를 살펴보면 키워드로 검색한 경우가 가장 많았으며 그 다음으로 서명, 등록번호, 저자명 등의 순서였다. 터미날의 위치별로는 자료실에 설치된 단말기에서 가장 많이 이용하였으며 자료실에서는 서명으로 검색한 경우가 가장 많고 그 다음으로 키워드, 등록번호의 순이고 연구부서에서는 키워드, 서명, 등록번호 순이며 일반부서에서는 저자명, 서명, 키워드 순이었다. 언어나 내용형식으로 검색한 경우와 set 와 검색명령을 조합한 경우는 상당히 적었다. 자료실에 설치된 터미날에서 검색회수가 많

은 이유는 일반 사용자들도 함께 사용하고 있기 때문이며, 서명으로 검색한 경우가 가장 많은 이유는 자료 구입시 복본조사를 주로 서명을 이용하여 수행하기 때문이다.

검색명령의 유형별로 살펴보면 등록번호와 분류기호는 자료실에서 거의 대부분(80% 이상)을 사용하였으며 두가지 이상의 검색 항목을 조합한 경우는 연구부서에서 가장 많이 나타났고 언어로 검색한 경우는 일반부서에 가장 많았다.

이는 자료실에서는 목록 작업등의 업무에 등록번호와 분류기호를 주로 사용하고 일반 이용자에게는 생소한 항목이기 때문이며 연구부서에서는 업무의 성격상 특정분야의 자료를 많이 찾는 경향이 있기 때문이다.

〈그림 1〉 검색명령 유형별 이용도

3.2 검색 실패 원인

검색 결과의 건수가 0 일때를 실패로 취급하였으며 앞에서 살펴 보았듯이 전체적으로 38.6%(5,713건)가 실패한 것으로 나타나 있다.

이를 원인별로 분석하면 〈표 2〉와 같으며 대부분 데이터베이스에 존재하지 않는 자료를 찾은 경우가 74.7%이다. 오타로 인하여 결과가 0으로 나온 경우가 16%였으며 사용상의 오류인 경우가 9.3%였다.

이를 좀더 상세하게 살펴보면 보유 장서수가 상대적으로 적은 소설류나 문제집을 찾는 경우가 있었고 주제어의 사용이 미숙하여 키워드와 같은 개념으로 사용한 것이 많았고 키워드 검색시 저자명을 key 로 사용한 경우도 많았다.

저자명으로 검색할 때는 외국의 저자명에 대한 사용법이 미숙하고 서명으로 검색할 때는 서명 전체를 입력함으로써 실패하는 경우와 자료실에서 복본조사를 위하여 검색한 경우가 많았다. 코드화 되어 있는 언어와 내용형식은 코드

〈표 1〉 사용자군에 따른 검색명령 유형별 이용도와 성공률

유형	순위	자 료 실			연 구 부 서			일 반 부 서			총 합					
		사용회수	사용비율	성공률	사용회수	사용비율	성공률	사용회수	사용비율	성공률	사용회수	사용비율	성공률			
등록번호	1,762	84.2	82.3	17.7	254	12.1	85.0	15.0	77	3.7	64.9	35.1	2,093	100	82.0	18.0
저자	667	32.8	56.7	43.3	202	10.0	46.5	53.5	1,161	57.2	47.4	52.6	2,030	100	50.3	49.7
분류기호	434	80.5	68.9	31.1	66	12.3	56.1	43.9	39	7.2	46.2	53.8	539	100	65.7	34.3
키워드	2,444	57.6	67.1	32.9	884	20.1	72.8	27.2	945	22.3	62.1	37.9	4,243	100	67.2	32.8
인어	16	13.8	75.0	25.0	31	2.7	51.6	48.4	69	59.5	66.7	33.3	116	100	63.8	36.2
출판년도	81	36.2	91.4	8.6	72	32.1	87.5	12.5	71	31.7	84.5	15.5	224	100	87.9	12.1
주제어	316	51.0	43.4	56.6	119	19.2	36.1	63.9	185	29.8	30.3	69.7	620	100	38.1	61.9
서경	2,773	67.2	52.1	47.9	350	8.5	56.3	43.7	1,005	24.3	42.7	57.3	4,128	100	50.2	49.8
내용형식	49	31.4	63.3	36.7	30	19.2	66.7	33.3	77	49.4	49.4	50.6	156	100	57.1	42.9
항목조합	128	54.9	60.2	39.8	39	16.8	61.5	38.5	66	28.3	57.6	42.4	233	100	59.7	40.3
SET 검색	14	26.9	92.3	7.7	35	67.3	71.4	28.6	3	5.8	33.3	66.7	52	100	74.5	25.5
SET 조합	169	46.6	87.0	13.0	107	29.5	81.3	18.7	87	24.0	72.4	27.6	363	100	81.8	18.2
계	8,853	59.8	64.4	35.6	2,159	14.6	66.9	33.1	3,785	25.6	51.1	48.9	14,797	100	61.4	38.6

비율단위 [%]

를 사용하지 않음으로써 실패하는 경우가 많았다. KTALIS는 화면에 사용가능한 모든 기능을 항상 보여 주기 때문에 사용자들이 지침서를 읽지 않고 사용하는 것도 검색 실패의 중요한 원인으로 생각된다.

현재의 여건에서 검색 성공율을 높이기 위하여 지속적인 사용자 교육이 필요하며 특히 코드화 되어있는 검색 대상 항목에 대한 교육이 잘 되어야 할 것으로 판단된다. 데이터베이스에 존재하지 않는 자료를 찾음으로써 검색에 실패하는 것을 줄이기 위해서는 당연히 보유장서를 확충하는 방법밖에 없으나 사용자가 많이 요구하는 자료가 어떤 것인가를 먼저 판단하여야 될 것이다.

<표2> 검색 실패의 원인

오 타	사용법미숙	자료없음	계
918 건 (16.0%)	529 건 (9.3%)	4,266 건 (74.7%)	5,713 건

3.3 단어 사용 빈도

<표3>은 키워드, 서명, 주제어 검색에서 많이 사용된 단어를 한글과 영문으로 분리하여 나열한 것이다. KTA의 업무와 관련이 큰 단어의 출현 빈도가 높았으며 전문성이 높은 단어보다 일반적인 단어들 이 많았다.

<표3> 사용빈도가 높은 단어

4. 시스템 응답시간

시스템 응답시간의 분석에는 검색결과 건수가 0이 아닌 9,084개의 레코드가 사용되었다.

이 레코드들에 기록되어 있는 응답시간과 CPU time을 분석하여 표와 그래프로 나타내었다. 그래프는 보기 쉽도록 log scale을 취

했으며 수평축은 검색결과 건수이고 수직축은 소요시간을 나타낸다. 소요시간이라 함은 응답시간이나 CPU time을 말하는데 응답시간은 사용자가 검색명령을 입력한 후 터미날에 그 결과를 통보 받기까지 걸린 전체 소요시간(CPU time, wait time, idle time)이고 CPU time은 검색명령의 수행을 위하여 사용된 순수시간(I/O time 포함)이다.

현재 KTALIS는 시스템의 전체 부하중 약11%를 차지하고 있으며 본고에서는 시스템 부하의 영향을 배제하기 위하여 CPU time을 위주로 분석하였으나 이해를 돕기 위하여 사용자가 실제로 대기하는 시간을 나타내는 응답시간 표도 동시에 분석하였다.

<그림2>는 응답시간을 기준으로 분석한 결과인데 컴퓨터 시스템 부하의 영향으로 결과건수대 응답시간의 뚜렷한 관계를 찾기가 힘들지만 대략 어느정도 이상의 결과 건수에 대하여 응답시간은 건수에 비례하는 것으로 볼 수 있다.

<그림3>은 CPU time을 기준으로 분석한 결과인데 CPU time은 컴퓨터 시스템의 부하에 영향을 받지 않는 수치이므로 응답시간을 기준으로한 <그림2>보다 명확하게 결과 건수대 소요시간의 관계를 보여주고 있다. <그림3>을 살펴보면 결과건수에 비해서 소요시간이 긴 그룹과 일정한 결과건수 이상에서 결과 건수와 소요시간이 대체로 비례관계를 나타내는 그룹으로 분리할 수 있다. 이를 규명하기 위하여 검색명령에 논리연산자나 관계연산자를 사용한 경우와 단일 검색항목에 대하여 단일 key로 검색한 경우를 분리하여 동일한 분석을 통하여 <그림4>와 <그림5>를 만들었다.

<그림4>는 단일항목을 단일 key로 검색한 경우(6,106 건)인데, 이 그래프를 살펴보면 결과건수가 적을 때는 결과건수에 대해서 소요시

간의 변화가 별로 없으나 결과건수가 늘어나면 소요시간은 결과건수에 거의 비례함을 알 수 있다.

〈그림 5〉는 검색명령에 연산자를 사용하여 여러가지 항목을 조합 검색하였거나 단일항목에 대하여 여러가지 key를 사용한 경우(2,978건)인데, 이 그래프와 〈그림 3〉, 〈그림 4〉를 비교하면 〈그림 3〉에 나타나 있는 결과건수에 비하여 소요시간이 긴 그룹은 거의 모두 〈그림 5〉에 나타나 있음을 알 수 있다. 따라서 간단한 검색명령보다 복잡한 검색명령이 보다 정확하게 요구하는 자료를 찾을수 있지만 소요시간이 길어짐을 알 수 있다.

이상의 결과를 종합하여 생각하면 KTALIS는 검색을 수행할때 inverted file을 찾는것보다 찾은 결과를 수집하고 조합하는데 더 많은 시간을 소모한다고 판단된다. 따라서 inverted file을 찾는 알고리즘이 $\log N$ 의 complexity이지만 결과 건수가 증가하면 소요시간은 결과건수에 비례하게 되는 것이다.

〈그림 6〉에서 〈그림 17〉까지는 앞에서 분류한 검색명령의 유형별로 결과 건수대 소요시간(CPU time)¹의 관계를 나타낸 것이다.

결과 건수대 응답시간²의 관계도 동일한 방법으로 분석하였으나 두 분석결과(1과 2)의 차이가 미소하여 본고에서는 결과 건수대 응답시간의 분석결과는 생략하였다.

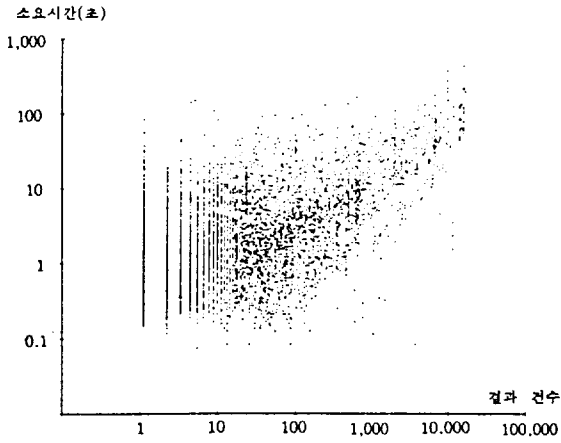
〈그림 6〉은 등록번호로 검색한 경우(1,716건)인데 결과 건수에 비해서 소요시간이 긴 부분은 주로 목록작성을 위하여 여러개의 등록번호를 연산자를 이용하여 조합한 경우이다. 또 결과 건수도 많고 소요시간도 긴 부분은 주로 관계연산자(greater than, less than 등)를 사용하여 검색한 경우이다. 〈그림 7〉은 저자

명으로 검색한 경우(1,022건)인데 대부분 1초 이내에 검색이 완료되었다. 〈그림 8〉은 분류기호로 검색한 경우(354건)인데 비교적 정상적인 분포를 보이고 있다. 〈그림 9〉는 키워드로 검색한 경우(2,850건)인데 결과 건수에 비해서 소요시간이 긴 부분은 대부분이 여러개의 key를 조합하고 검색한 경우이다. 〈그림 10〉은 언어로 검색한 경우(74건)인데 결과 건수가 많고 소요시간이 긴 부분이 많았다.

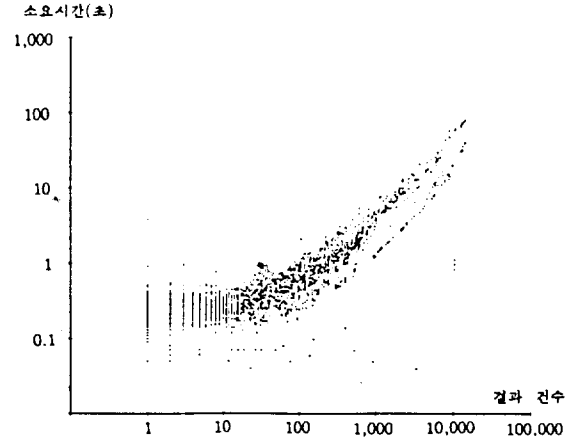
이것은 현재 보유하고 있는 자료의 대부분이 한글, 영어, 일본어로 되어 있기 때문에 당연한 결과로 볼 수 있다. 〈그림 11〉은 출판년도로 검색한 경우(197건)인데 대체로 결과 건수와 소요시간 모두가 큰편이다. 〈그림 12〉는 주제어로 검색한 경우(236건)인데 비교적 정상적인 분포를 보이고 있다. 〈그림 13〉은 서명으로 검색한 경우(2,071건)인데 비교적 정상적인 분포를 보이고 있다. 〈그림 14〉는 내용형식으로 검색한 경우(89건)인데 대체로 결과 건수와 소요시간 모두가 큰편이다. 〈그림 15〉는 둘이상의 검색항목을 조합하여 검색한 경우(139건)인데 대체로 소요시간이 긴편이고 조합하는 항목에 따라서 결과 건수와 소요시간에 큰 차이가 있다.

〈그림 16〉은 검색 결과로 만들어진 set와 검색명령을 조합한 경우(39건)인데 결과 건수는 대체로 적은 편이나 소요시간은 조합 방법에 따라서 큰차이를 보이고 있다. 〈그림 16〉은 검색 결과로 만들어진 set끼리 조합한 경우(297건)인데 대체로 결과 건수에 상관없이 소요시간이 짧은 분포이다.

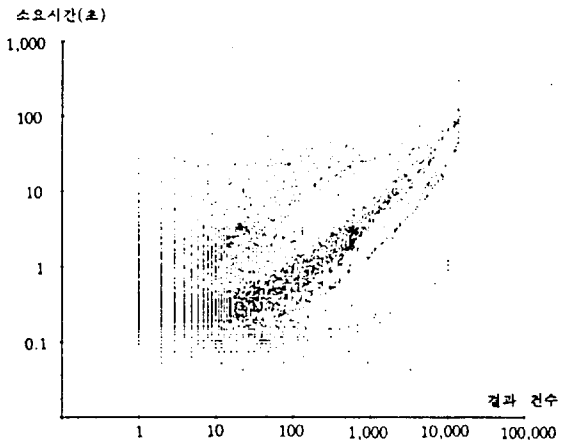
〈표 4〉는 검색명령의 유형별로 일정 시간내에 검색이 완료되는 비율을 나타내고 있는데 CPU time이 10초 미만인 검색명령이 97%



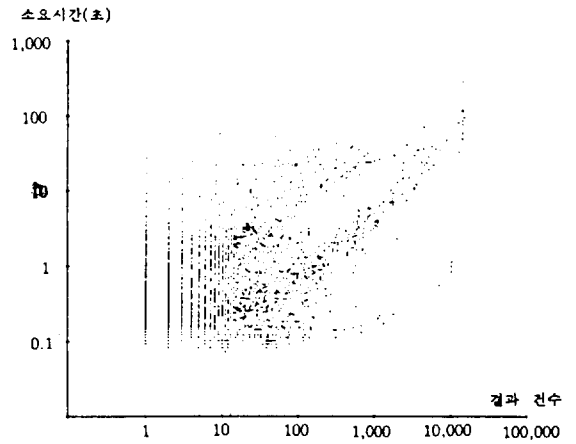
<그림 2> 결과건수 대 응답시간



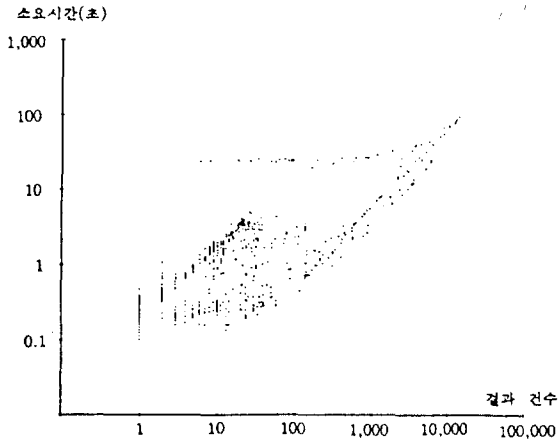
<그림 4> 단일항목 단일검색키 사용



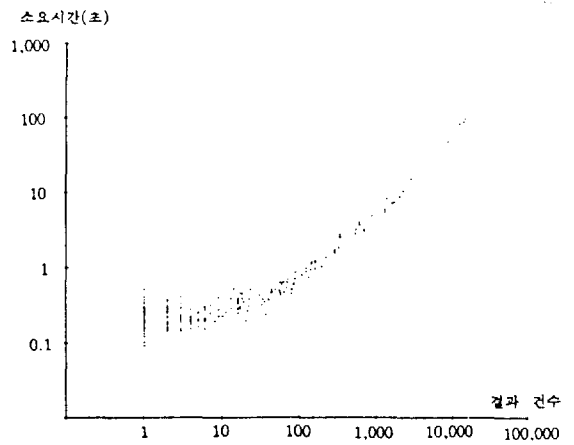
<그림 3> 결과건수 대 CPU time



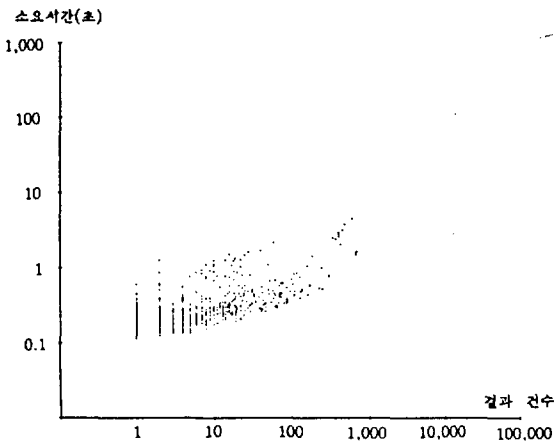
<그림 5> 연산자 사용 조합검색



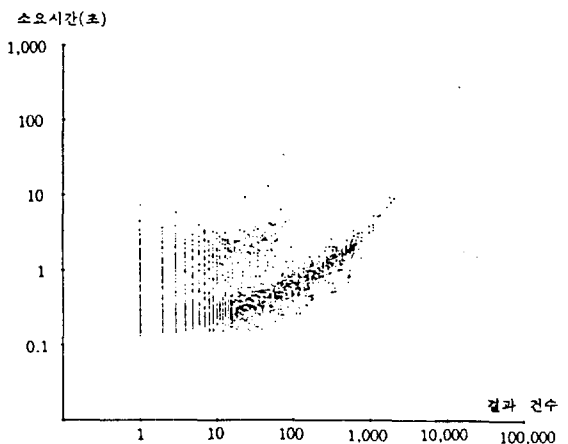
<그림 6> 등록번호 검색



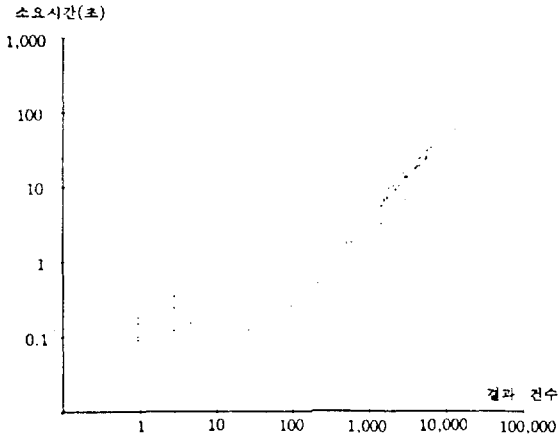
<그림 8> 분류기호 검색



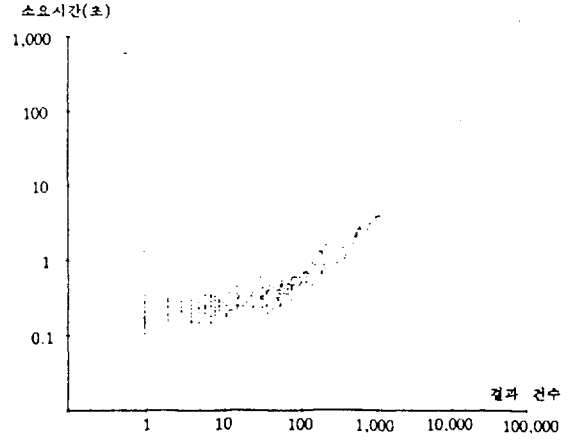
<그림 7> 저자명 검색



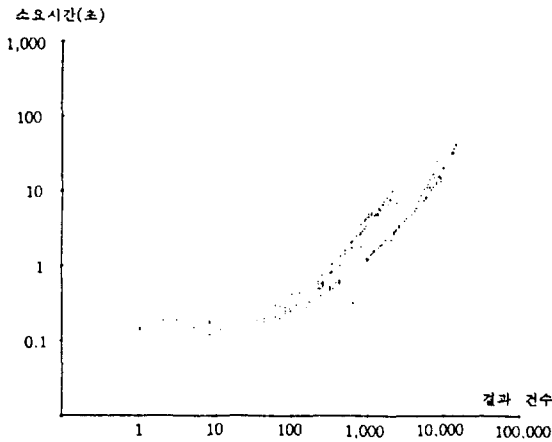
<그림 9> 키워드 검색



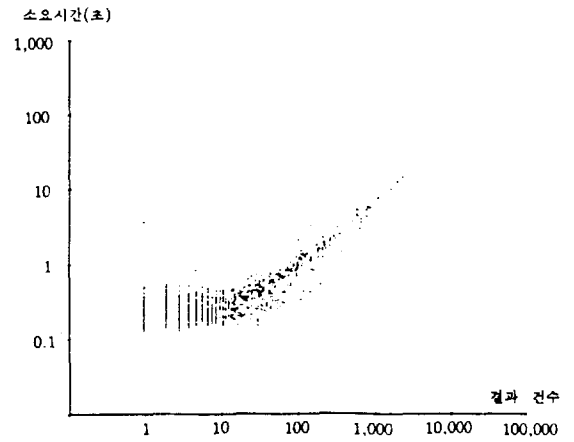
<그림 10> 언어 검색



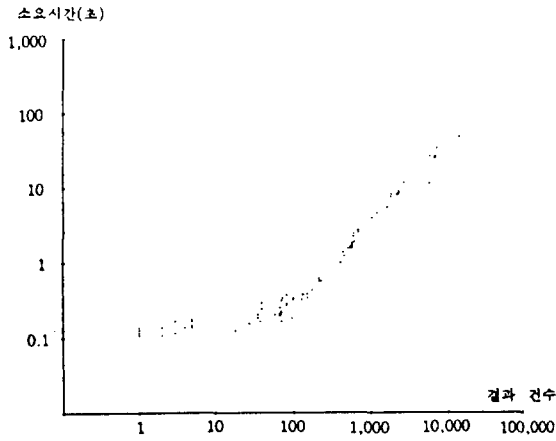
<그림 12> 주제어 검색



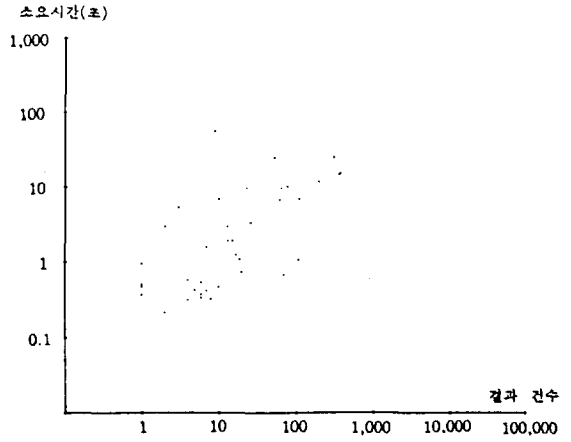
<그림 11> 출판년도 검색



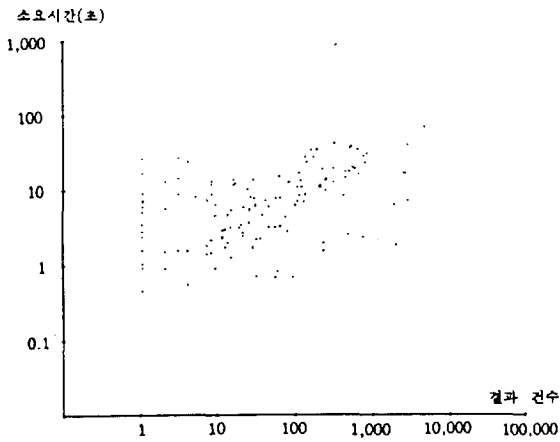
<그림 13> 서명 검색



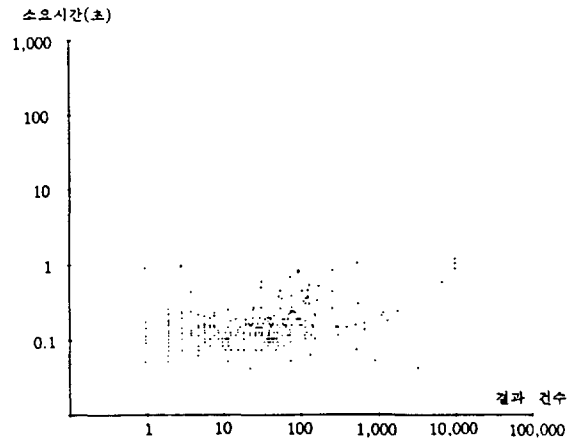
<그림 14> 내용형식 검색



<그림 16> SET 검색



<그림 15> 항목 조합 검색



<그림 17> SET 조합

를 차지하고 있으며 CPU time이 1초 미만인 것은 81.2%를 차지하고 있다. 응답시간은 1초 미만이 42.7%, 1초에서 10초 사이는 49.2%로 10초 미만이 91.9%를 차지하고 있다. 언어, 출판년도, 내용형식을 검색한 경우에 평균 결과 건수가 많으며 소요시간이 긴 것이 많은데 언어의 경우 CPU time으로는 10초에서 30초사이가 52.7%로 가장 많고 응답시간으로는 30초에서 60초 사이가 32.4%로 가장 많다. 또 응답시간이 60초 이상인 것도 20.3%나 되었다.

< 표 4 >

< 표 5 >는 검색명령의 유형별로 일정비율의 검색이 완료되는 시간을 응답시간과 CPU time으로 나타내고 있다. 이 표를 보면 전체적으로 50%의 검색이 완료되는데-즉 10개의 검색명령중 5개가 완료-CPU time으로는 0.27초, 응답시간으로는 1.23초가 소요되었다. 90%의 경우 CPU time 2.09초, 응답시간 8.23초가 소요되었고 최악(100%)의 경우 CPU time 285.03초, 응답시간 415.7초가 소요되었다. 이 표에서도 언어로 검색한 경우의 소요시간이 가장 긴 것으로 나타나고 있다.

< 표 5 >

5. 결 론

이상에서 KTALIS의 개통 초기부터 검색시스템의 LOG file에 저장된 검색결과와 응답시간에 관한 데이터를 여러가지 각도에서 분석하여 표와 그림으로 작성하여 보았다.

일반적으로 사용자들은 키워드, 서명, 등록번호, 저자명을 검색 대상으로 많이 선택하였으며 등록번호와 분류기호는 자료실에서 거의 대부분을 사용하였고 둘이상의 검색항목을 조합한 경

우는 연구부서에서 가장 많았다.

검색 성공율은 전체적으로 61.4%였으며 출판년도로 검색한 경우가 성공율 87.9%로 가장 높았고 주제어로 검색한 경우의 성공율이 38.1%로 가장 낮았다. 검색 실패의 원인은 주로 데이터베이스에 존재하지 않는 자료를 검색한 경우이며 오타나 검색시스템에 대한 이해부족인 경우도 많았다.

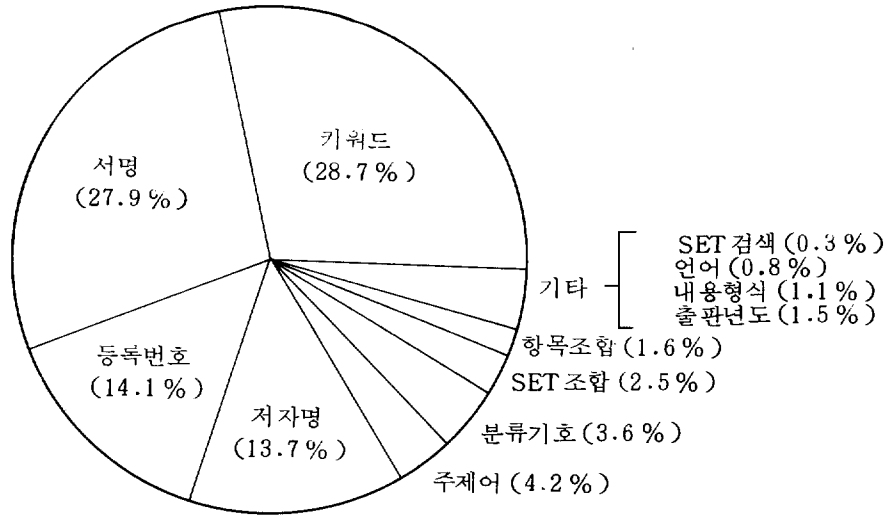
시스템의 응답시간은 대략 결과 건수에 비례하지만 결과 건수가 적을 때는 비슷한 응답시간을 보였다. 코드화되어 있는 출판년도, 언어 내용형식으로 검색한 경우 대체로 많은 시간이 소요되었다. 시스템 전체적으로는 90%의 검색명령이 CPU time 2.09초이내, 응답시간 8.23초이내에 완료되었으며 실제로 사용자의 대기시간에 대한 불만도는 매우 적은 것으로 조사되었다. 이것은 15-20초정도를 사용자의 대기시간 불만족 기준치로 삼고 있는것 [2]에 비하면 중형 컴퓨터로서 도서관 전용의 시스템이 아닌 점(시스템 점유율 약 11%)을 감안하면 매우 양호한 시스템의 성능으로 판단된다.

이 결과는 KTALIS의 개통후 1년 정도의 데이터를 토대로한 것으로 향후 KTALIS를 계속 추가개발해 나가는데 중요한 자료로 사용될 것이며 타 기관에서 정보검색시스템을 설계하는데 유용한 자료가 될 수 있을 것이다.

* 참고문헌

1. Kunj B. Rastogi, Ichiko T. Morita, "OCLC Search Key Usage Patterns in a Large Research Library", Journal of Library Automation Vol14/2 1981, pp.90-99

2. Walt Crawford, "Long Searches, Slow Response : Resent Experience on RLIN", Information Technology and Libraries Jun. 1983, pp.176-182
3. Kunj B. Rastogi, "Maximun and Minimun Derived Search Key Use and Its Relationship with Found and Not-found System Responses," Journal of Library Automation Vol112/2, 1979, pp.169-174
4. 김재일, 김영현, 이해옥, "AVL 트리를 이용한 Inverted 화일구조 구현과 도서 데이터베이스," 한국정보과학회 '87 봄 학술 발표논문집, pp.327-330
5. 최상규, 허문생, "도서정보관리시스템 개발 실용화," 한국전기통신, 1987, pp.22-27.
6. 이해옥, 김영현, 김재일, 허문행, "정보관리 토달시스템 구축을 위한 KTALIS개발," 정보관리 학회지 4 권 2 호, 1987, pp.152-172
7. AT&T, "UNIX System V Programmer Reference Manual," 1984
8. AT&T, "UNIX System V User Reference Manual," 1984
9. B.W.Kernighan, D.M.Ritchie, "The C Programming Language," Prentice-Hall, 1978



<그림 1> 검색명령 유형별 이용도

한	글	영	문
통신	관리	computer	international
경영	계획	data	power
컴퓨터	케이블	telecommunication(s)	electrical
전자	경제	software	handbook
전기통신	감사	system (s)	theory
정보	구조	management	programming
전기	국제	engineering	technology
통계	경영학	communication (s)	processing
기술	최신	information	language
전자회로	공학	control	LAN
전자공학	전자기학	analysis	CCITT
교육	방법	network (s)	research
회계	인사	quality	guide
시스템	네트워크	digital	business
정보통신	행정	design	public
한국	전화	NTT	IBM
기업	설계	database	2000
분석	데이터	reliability	TDX
기준	회로	market(ing)	signal
정책	정보처리	ISDN	artificial
사전	전산화	C	world
투자	영어	electronic (s)	statistics
조사	보고서	IEEE	standard
마케팅	기초	CATV	security
마이크로	규격	telephone	VLSI

<표 3> 사용빈도가 높은 단어

〈표 4〉 검색명령 유형별 일정시간 내의 검색완료 비율

구 유형	시간 사용회수	1 초이내		1 ~ 10 초		10 ~ 30 초		30 ~ 1 분		1 분이상	
		CPU	TOTAL	CPU	TOTAL	CPU	TOTAL	CPU	TOTAL	CPU	TOTAL
등록 번호	1,716	72.7	54.5	21.4	35.6	4.7	5.7	0.7	2.7	0.5	1.6
저자 명	1,022	96.7	50.4	3.2	47.9	0.0	1.5	0.0	0.1	0.1	0.1
분류 기호	354	90.4	60.2	7.9	33.6	0.6	4.5	0.3	0.6	0.8	1.1
키워드	2,850	76.3	25.2	23.6	67.7	0.0	6.6	0.0	0.4	0.0	0.1
인어	74	14.9	12.2	23.0	5.4	52.7	29.7	9.5	32.4	0.0	20.3
출판년도	197	31.5	16.8	49.7	43.1	10.7	22.3	8.1	12.7	0.0	5.1
주제어	236	86.9	33.1	12.3	63.1	0.4	3.4	0.0	0.0	0.4	0.4
서명	2,071	96.5	53.3	3.3	45.0	0.1	1.4	0.0	0.3	0.1	0.1
내용형식	89	60.7	36.0	29.2	40.4	6.7	12.4	3.4	5.6	0.0	5.6
항목조합	139	6.5	0.7	54.7	27.3	32.4	46.8	5.8	15.8	0.7	9.4
SET 검색	39	42.1	7.9	39.5	52.6	15.8	31.6	2.6	5.3	0.0	2.6
SET 조합	297	99.0	80.1	1.0	19.5	0.0	0.3	0.0	0.0	0.0	0.0
총합	9,084	81.2	42.7	15.8	49.2	2.3	5.6	0.6	1.6	0.2	0.9

* 단위는 (%) 임

<표 5> 검색명령 유형별 일정비율 검색완료 시간

구 비율 유형	50 %			70 %			90 %			Worst Case		
	결과건수	CPU	TOTAL	결과건수	CPU	TOTAL	결과건수	CPU	TOTAL	결과건수	CPU	TOTAL
등록 번호	1	0.21	0.80	9	0.66	2.26	89	3.24	9.81	13,954	83.31	352.91
저자 명	3	0.22	0.99	7	0.27	1.61	28	0.47	3.40	13,954	114.68	209.30
분류기호	2	0.20	0.73	5	0.28	1.38	135	0.95	4.10	14,520	94.63	262.49
키워드	16	0.39	1.91	48	0.78	3.53	248	1.86	8.51	13,952	285.03	415.70
언어	3,450	13.49	33.02	5,988	23.87	44.38	6,086	25.52	74.84	13,955	58.69	168.44
출판년도	1,091	2.93	5.69	2,142	5.65	13.91	8,615	16.19	44.23	13,955	42.39	123.96
주제어	23	0.31	1.50	60	0.46	2.53	387	1.47	6.84	13,955	76.30	101.79
서명	4	0.24	0.94	11	0.31	1.51	40	0.51	3.47	13,954	118.87	200.28
내용형식	128	0.35	1.71	571	1.75	4.38	2,238	9.31	35.97	13,955	49.19	117.21
항목조합	30	7.28	15.30	126	11.73	26.93	470	23.97	54.30	4,385	70.12	156.26
SET 검색	10	1.23	7.05	26	6.81	11.94	109	15.03	25.24	1,578	55.85	76.75
SET 조합	27	0.14	0.33	52	0.17	0.61	147	0.32	2.01	10,310	1.14	15.09
총합	7	0.27	1.23	24	0.46	2.48	196	2.09	8.23	14,520	285.03	415.70

* CPU, TOTAL time의 단위는 [초]

참 고 문 헌

1. Kunj B. Rastogi, Ichiko T. Morita, "OCLC Search Key Usage Patterns in a Large Research Library," *Journal of Library Automation* Vol. 14/2, 1981, pp. 90-99.
2. Walt Crawford, "Long Searches, Slow Response: Resent Experience on RLIN," *Information Technology and Libraries* Jun. 1983, pp. 176-182.
3. Kunj B. Rastogi, "Maximum and Minimum Derived Search Key Use and Its Relationship with Found and Not-found System Responses," *Journal of Library Automation* Vol. 12/2, 1979, pp. 169-174.
4. 김재일, 김영현, 이해옥, "AVL 트리를 이용한 Inverted 화일구조 구현과 도서데이터베이스" 한국정보과학회 '87봄 학술발표논문집, pp. 327-330.
5. 최상규, 허문행, "도서정보관리시스템 개발 실용화," 한국전기통신, 1987, pp. 22-27.
6. 이해옥, 김영현, 김재일, 허문행, "정보관리 토달 시스템 구축을 위한 KTALIS 개발," 정보관리학회 지 4권 2호, 1987, pp. 152-172.
7. AT&T, "UNIX System V Programmer Reference Manual," 1984. 1984.
8. AT&T, "UNIX System V User Reference Manual," 1984.
9. B. W. Kernighan, D.M. Ritchie, "The C Programming Language," Prentice-Hall, 1978.