

전화번호화일 관리시스템 개발

김의경 강종수 김영곤 이종원

한국통신 연구개발단 생활정보서비스연구실

Development of File System for Telephone Subscriber

Eui-Kyoung Kim, Jong-Soo Kang, Young-Gon Kim, Jong-Won Rhee

Electric Directory Service Lab, Korea Telecom Research Center

요 약

전자전화번호부 서비스는 HiTEL 단말기 및 기존의 PC를 소유한 가입자의 전화번호 관련문의에 대해 신속, 정확한 정보를 제공하는 비디오텍스 응용 시스템으로서 통신처리시스템, 서비스검색시스템, 화일관리시스템으로 나뉠 수 있다. 전화번호화일은 가입자에 따라 인명부, 업종상호부로 나누어 CISAM 방식의 화일로 구현하였으며, 상호인 경우 1:N 필드관계를 지원하여 하나의 전화번호에 대해 다수 업종의 지원이 가능하다. 전화가입자의 일일 변동자료의 처리 및 화일의 갱신에 대해 소개하고자 한다.

I. 서론

현재 한국통신공사에서는 가입자의 전화번호 관련정보를 고객관리시스템(TSIS : Telephone Subscriber Information System, 이하 TSIS라 지칭)과 신형114안내 시스템(KDAS : KT Directory Assistance System, 이하 KDAS라 지칭)에서 보유, 운영하고 있다. 그러나 전화가입자의 전화번호 관련정보가 TSIS에서는 요금위주의 업무를 토대로 하는 화일시스템이기에 안내업무의 성격에는 부적합 하다. 또한 KDAS는 광고화면이 제공되지 않는 ASCII위주의 문자형 전화번호 관련정보의 114안내요원이라는 중개자를 통한 안내업무이다. 또한 KDAS는 업종코드의 효율적인 관리가 수행되지 않기에 ASCII형태의 정보뿐만 아니라 NAPLPS(North American Presentation Level Protocol Syntax) 방식의 광고화면이 제공되는 그림형 전화번호 관련 정보를 중개자의 도움 없이 가입자가 직접 검색하는 전자전화번호부서비스(EDS : Electronic Directory Service, 이하 EDS라 지칭)에는 부적합하다. 그러므로 EDS에서는 그림형 전화번호 관련정보를 가입자가 직접 검색하는 업무성격을 위해 고유의 전화번호 관련정보 화일시스템을 보유, 관리하여야 한다[1,2,3].

EDS 전화번호 화일관리 시스템은 KDAS로부터 가입자의 전화번호 관련정보를 제공받아 색인화된 순차적 접근방식(CISAM : C Indexed Sequential Access Method, 이하 CISAM이라 지칭)을 제공하여 다량의 데이터를 처리, 신속한 검색을 지원하는 CTL(비디오텍스 S/W 상품명) 화일 관리 특성에 적합한 형태로 변환하여 EDS 전화번호 초기 화일을 구축한다. 자료처리센터에서 매일 발생하는 전화가입자의 변동사항을 TSIS와 KDAS에서 수집하여 EDS 화일 특성에 적합한 일일변동자료를 생성하여 EDS 전화번호화일을 현행화 한다.

EDS 화일관리 시스템은 매우 많은양의 가입자 전화번호 관련정보를 처리할뿐만 아니라 신속한 검색속도를 보장하기 위해 B-tree기법 CISAM 방식의 CTL화일 관리 모듈을 응용하여 개발하였다. 이에 II장에서는 CTL화일 관리모듈을 소개하며, III장에서는 CTL화일 관리모듈을 응용한 EDS 화일 관리시스템의 개발 내용을 정리하였다.

II. CTL 화일관리모듈 소개

아래와 같은 특성을 갖는 m 원 B-tree기법을 이용 색인화일을 구성하는 CISAM을 기본적으로 제공하는 CTL 화일관리 모듈은 UNIX환경하에서 다량의 데이터를 처리, 신속한 검색을 지원한다.[8,9]

- 루트(root)와 단말(leaf)을 제외한 모든 노드는 최소 $\lceil m/2 \rceil$, 최대 m 개의 서브트리를 갖는다.
- 루트는 단말이 아닌 이상 적어도 두 개의 서브트리를 갖는다.
- 모든 단말노드는 같은 레벨에 있다.
- 단말이 아닌 노드의 키값 수는 그 노드의 서브트리수보다 하나 적으며, 각 단말 노드는 최소 $\lceil m/2 \rceil - 1$ 개, 최대 $m - 1$ 개의 키값을 갖는다.
- 한 노드안에 있는 키값들은 오름차순을 유지한다.

본장에서는 CTL화일 관리모듈에서 기본적으로 구성되는 화일구조와 화일 관리상의 특성 및 제공하는 유틸리티(utility)를 간략히 소개하였다.

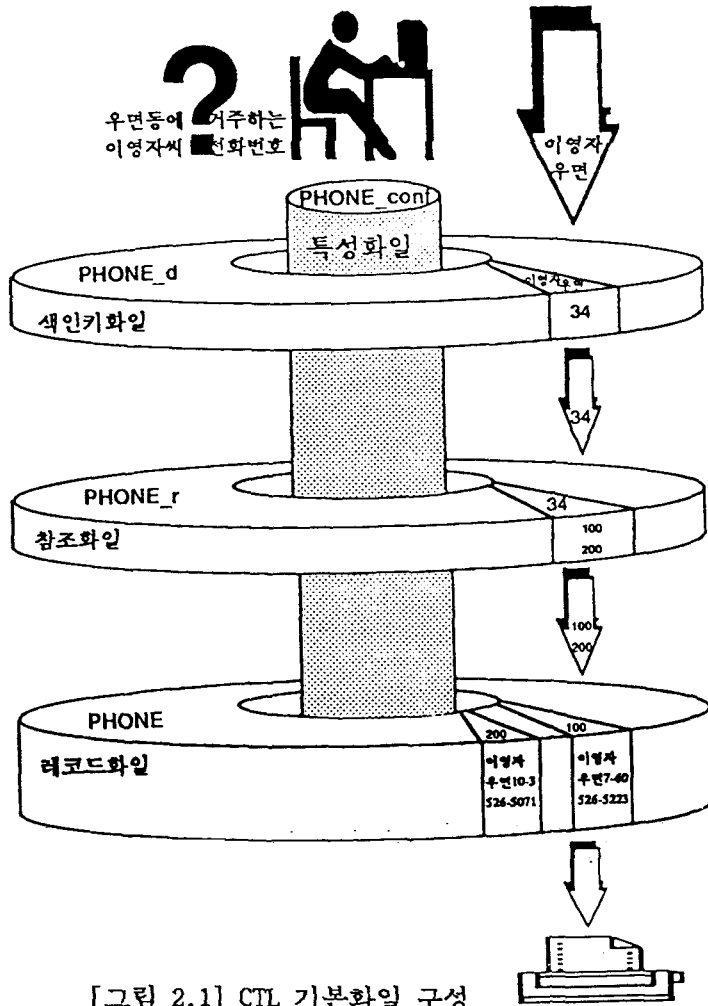
1. 화일구조

CISAM 구조의 CTL화일 관리모듈은 기본적으로 키정의에 따라 B-tree 구조상에 키값이 색인화된 색인키화일, 검색키 색인에 따라 해당 레코드의 주소가 수록되어 있는 참조화일, 실제 데이터 레코드가 저장되어 있는 레코드화일 및 키와 레코드 구성의 정보가 수록되어 있는 특성화일의 4개 기본화일을 필요로 한다.

CTL화일 구조는 색인키화일에 해당 레코드 주소를 지정하지 않고 독립적으로 참조화일에 지정함으로써 정의된 키값에 대해 다수의 레코드 관리면에서 색인키화일상의 B-tree 노드 메모리 낭비를 줄일 수 있으며, 색인키화일에서 검색키에 의한 입/출력 노드정보의 양을 줄일 수 있어 검색속도를 줄일 수 있다.

[그림 2.1]은 검색키에 대한 해당 레코드의 검색과정면에서 CTL 구성화일을 나타낸 것으로 검색자가 “우면동에 거주하는 이영자” 전화가입자의 전화번호를 검색코자할 때, CTL 구성화일간의 연동관계를 나타낸 것이다. 가입자 이름과 주소의 동명을 키로 정의한 “PHONE”이라는 화일시스템에서 검색자가 가입자 이름 “이영자”와 동명 “우면”을 입력하였을 경우 CTL화일 관리모듈은 아래와 같은 작업을 순차적으로 수행하여 해당레코드를 검색하여 준다.

- (1) 주어진 검색키(“이영자 우면”)를 특성화일(PHONE_conf)에서 정의된 키구성의 적합성을 검증한다.
- (2) 색인키화일(PHONE_d)에서 검색키의 해당 노드를 탐색하여 참조주소(“34”)를 발췌한다.
- (3) 참조주소로 참조화일(PHONE_r)에서 해당하는 2개의 레코드주소(“100”, “200”)를 발췌한다.
- (4) 레코드주소로 레코드화일(PHONE)에서 해당 2개의 레코드(“이영자:우면10-3:526-5071:”, “이영자:우면7-60:526-5223:”)를 검색한다.



[그림 2.1] CTL 기본파일 구성

2. 파일관리특성

CTL파일 관리모듈은 고정길이(fixed length) 레코드형태에서 단순키, 복합키 및 임의키로 키를 정의할 수 있으며, 정의된 키값의 전체 혹은 앞부분으로서 레코드 검색이 가능하며 다중사용자를 지원하기 위해 공유메모리(shared memory)기법을 제공한다.

가. 레코드 형식

CTL의 데이터레코드는 구분자(delimiter) ':'를 이용한 가변길이 필드 구조로서 내부적 낭비(internal fragmentation)는 발생하지만 고정길이 레코드 구조로서 외부적 낭비(external fragmentation) 요소는 없어, 레코드내용이 자주 변경되는 파일시스템에서는 효율적으로 메모리를 관리할 수 있다. KSC5601 2바이트 완성형 한글, 대문자 영문, 숫자 및 특수 기호의 자료표현(cannonical form)을 갖으며 아래와 같은 형식을 취한다.[7]

- 레코드상의 각 필드는 ':'로 구별되어 진다.
- 레코드의 길이는 최소 32 byte이상이어야 한다.
- 모든 레코드는 동일 레코드 길이를 취한다.
- 고정길이 레코드를 유지하기 위해 space 글자를 삽입한다.
- 레코드내용의 마지막 글자는 CR(carriage return)이어야 한다.

나. 키구성 및 검색방법

CTL화일 관리모듈에서는 하나 이상의 필드를 이용하여 키구성이 가능하며 사용되는 각 필드는 전내용 혹은 일부분 내용으로 키구성이 된다. 또한 정의된 키값의 전체 혹은 앞부분의 내용으로 해당 레코드의 검색이 가능하다.

1) 키구성

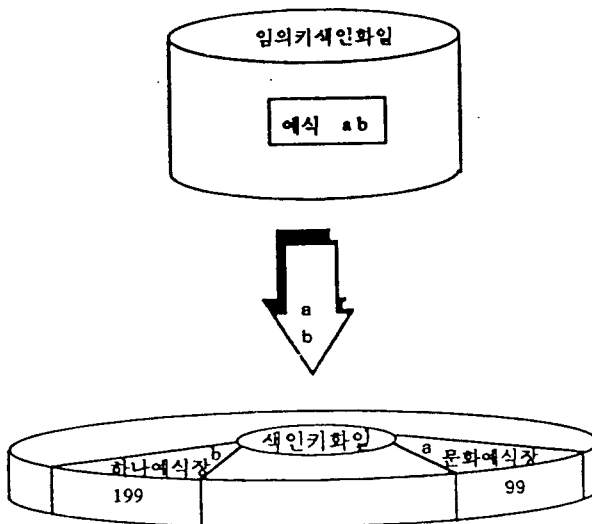
CTL이 제공하는 키구성 방법은 아래와 같다.

- 단순키(simple key) : 한개의 필드로 키구성.
- 복합키(composite key) : 복수개 필드로 키구성.
- 임의키(virtual key) : 단순키로서 필드내용상의 임의의 부분 내용으로 키구성.

단순키와 복합키에서 사용되는 필드는 필드내용 전부 혹은 앞부분의 내용으로 키를 정의할 수 있으며, 한개의 필드로 구성되는 임의키는 필드의 임의 부분내용으로 키구성이 가능하도록 키구성전에 가능한 필드부분내용으로 검색키값의 정의영역(domain)을 텍스트형태의 화일에 수록한다. 임의키 검색을 위한 임의키 정의영역을 색인화한 임의키색인 화일을 별도로 생성하여 기존의 색인키화일과 연동할 수 있도록 한다.

2) 검색방법

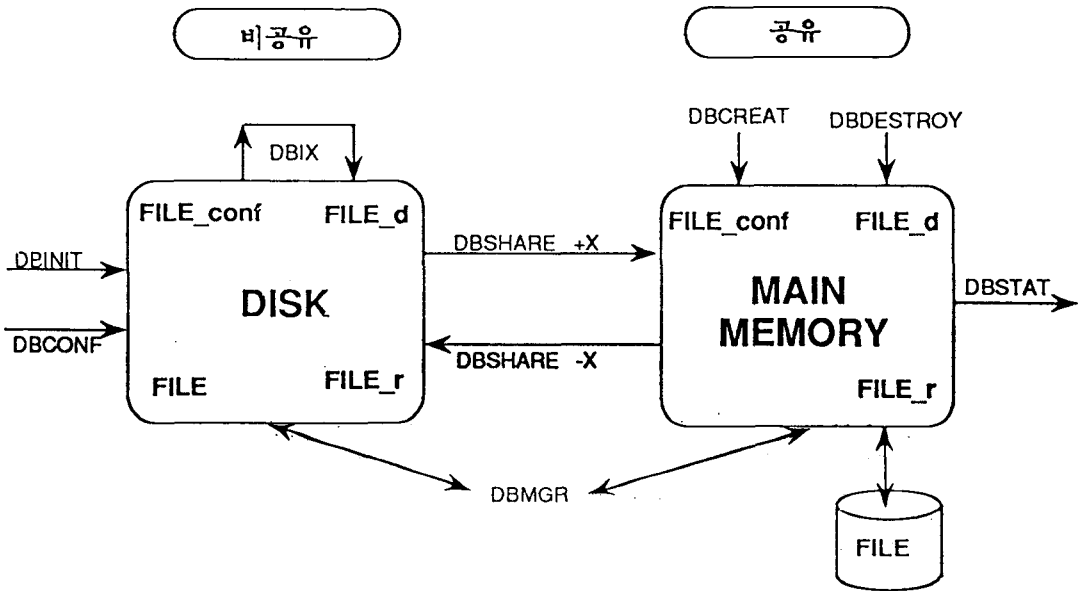
정의된 키값의 전체 혹은 앞부분의 내용으로 해당레코드의 검색이 가능하며 임의키 검색과정은 [그림 2.2]와 같다. 임의검색키 정의영역 화일상의 내용으로 색인화된 별도의 임의키색인화일이 존재하여 기존의 색인키화일과 연동하여 해당레코드를 검색한다. [그림 2.2]에서는 색인키화일까지의 연동과정만 표현하였기에 해당레코드 검색까지의 과정은 [그림 2.1]을 참조하기 바란다.



[그림 2.2] 임의키 검색과정

다. 다중사용자 환경

여러 사용자가 화일시스템을 공유함으로써 동시에 실행되는 트랜잭션(transaction)들이 화일의 일관성(consistency)을 유지하기 위해 CTL화일 관리 모듈은 UNIX환경하의 공유메모리(shared memory)기법을 제공한다. 화일관리자는 하드디스크상이 아닌 주기억장치(main memory)상에 화일시스템의 구성화일을 로드(load)하여 여러 사용자가 동시에 화일시스템을 공유할 수 있도록하며, 또한 공유된 화일시스템의 현재 공유상황을 감지할 수 있는 기능도 제공한다. [그림 2.3]은 공유와 비공유 화일시스템의 구조를 나타낸 것이다.



[그림 2.3] 공유화일시스템의 구조

3. 제공기능

CTL화일 관리모듈은 화일시스템의 생성및 삽입, 삭제, 검색등의 기능을 UNIX환경하의 명령어(command)형태로 제공하며, 또한 C 프로그래밍상에서 이용할 수 있는 라이브러리(library)형태도 제공한다.

III. EDS 화일시스템 개발

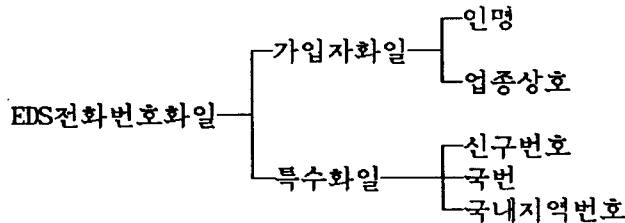
가입자 전화번호 관련정보를 컴퓨터의 디스크에 저장하는 화일조직은 저장된 레코드의 접근 방법에 따라 순차(sequence), 인덱스(index) 및 해싱(hashing) 방법으로 나눌 수 있다. 순차방법은 다수의 가입자에 대해 일정 순서로 레코드를 유지하며, 검색코자하는 레코드 이전의 레코드를 접근해야 하기에 검색속도면에서 EDS 화일 관리시스템에는 부적합하다. 해싱방법은 전화번호 관련정보상의 키값과 레코드 주소간의 해싱함수가 단순하지 않음으로 EDS에는 비효율적인 방법이 된다. 그러므로 EDS 전화번호 관련정보 화일 관리시스템은 아래와 같은 특성때문에 CISAM을 채택하였다.

- 다량의 데이터
서울지역인 경우 350만의 인명가입자와 60만 단체가입자에 대한 전화번호 관련 정보를 생성, 현행화 및 검색 등의 관리가 필요하다.
- 검색속도
B-tree 기법은 최적의 방법으로 검색자가 원하는 정보를 신속히 검색될 수 있다.
- 데이터처리의 단순성
화면설계, 보고서, 메뉴처리기능 등의 부가기능이 EDS상에는 필요치 않고 단순히 원하는 정보의 검색속도와 현행화기능의 단순한 데이터 처리가 요구된다.

본 장에서는 EDS의 전화번호정보의 구성화일과 초기화일의 구축과정, 가입자의 전화번호 변동사항 정보수집 및 구성된 화일에 변동사항을 적용하는 현행화 과정에 대해서 살펴보자.

1. 화일구성

EDS상의 가입자 전화번호 관련정보는 인명, 업종상호의 가입자화일과 부가 서비스를 위한 신규번호, 국번, 국내지역번호의 특수화일로 [그림 3.1]과 같이 나눌 수 있다.



[그림 3.1] EDS 전화번호 관련화일

각 화일의 정의는 아래와 같으며 화일특성은 <표 3.1>과 같다.

- 인명(RF : Residence File)
전화번호 정보의 80%를 차지하는 인명가입자의 정보
- 업종상호(YPBF : Yellow Page Business File)
업종코드가 수록된 기관, 상호단체가입자의 정보
- 신규번호(CH : Changed Number File)
가입자의 전화번호 변동내용의 정보
- 국번화일(PFX : Prefix Number File)
전화번호상의 국번에 대한 해당 영업국과 수용국 정보
- 국내지역번호화일(LAC : Local Area Code File)
시의전화 사용시 해당지역의 DDD번호 정보

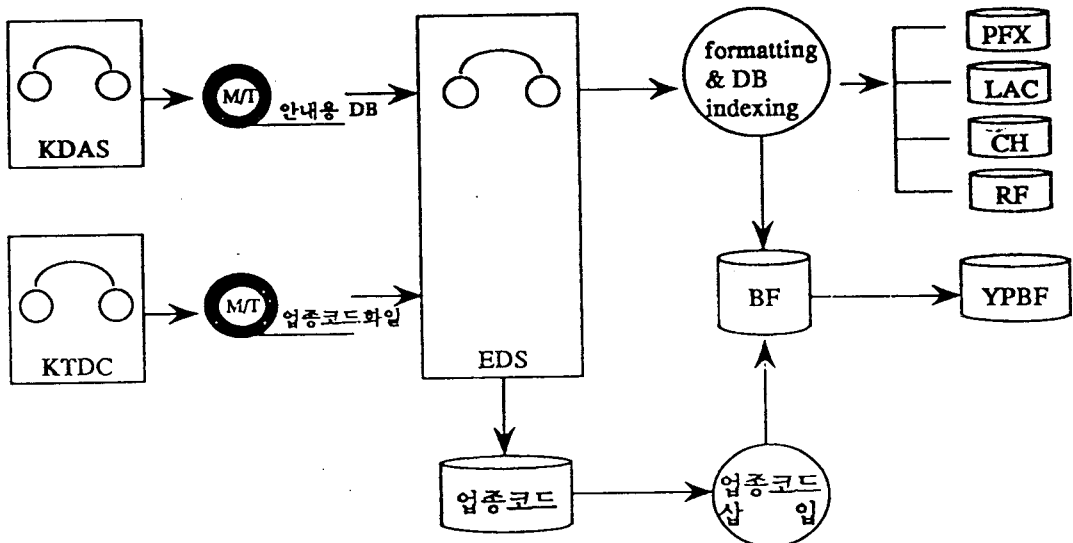
<표 3.1> EDS전화정보화일 내역

화일명	레코드 길이	레코드 갯 수	필드	키구성
인명	65	350만	이름, 주소, 전화번호, 게재상태*1	이름+주소+게재상태 전화번호
업종상호	300	55만	상호명, 동명, com*2, 전화번호, 주소, 업종코드, 게재상태*1, 부서명4개, 업종가능*3, 광고관련정보*4	상호명+동명+부서명 전화번호 업종코드
신구번호	32	16만	가입자명, 변경전전화번호, 변경후전화번호, 변경일자	변경전전화번호 변경후전화번호 변경일자
국번	32	600	국번, 영업국명, 수용국명	국번
국내지역	32	1400	도시명, 해당도명, DDD번호	도시명

- *1 : 가입자의 해당 전화번호 검색가능 유무
- *2 : 행정상 한동의 동명회사 구분자
- *3 : 업종서비스시 검색레코드 지정
- *4 : 광고화면의 특성및 화면이름

2. 초기화일 구축

KDAS의 안내용 화일을 순차화일로 변환, M/T를 이용하여 EDS 가입자 전화번호 관련정보 화일을 수집하여 CTL화일 관리모듈의 레코드화일 특성에 적합한 형태로 변환하여 [그림 3.2]와 같이 색인화하여 CISAM 형태의 전화번호 관련정보 화일을 초기 구축한다.



[그림 3.2] EDS 전화번호화일 초기구축 과정

EDS 화일상의 레코드는 아래와 같은 특성을 소유한다.

- 인명인 경우 게재불요가입자의 검색불능
- 업종상호에서의 한 전화번호에 대한 상하관계 부서 4개까지 수록가능
- 업종상호에서의 한 전화번호에 대한 복수개 업종코드 수록 가능
- 검색자위주의 상호명처리 가능
- 광고화면 정보 수록
- 업종서비스시에 업종 flag가 존재하여 업종에 대한 해당 단체의 소수 레코드 검색 가능.

3. 일일변동자료 생성

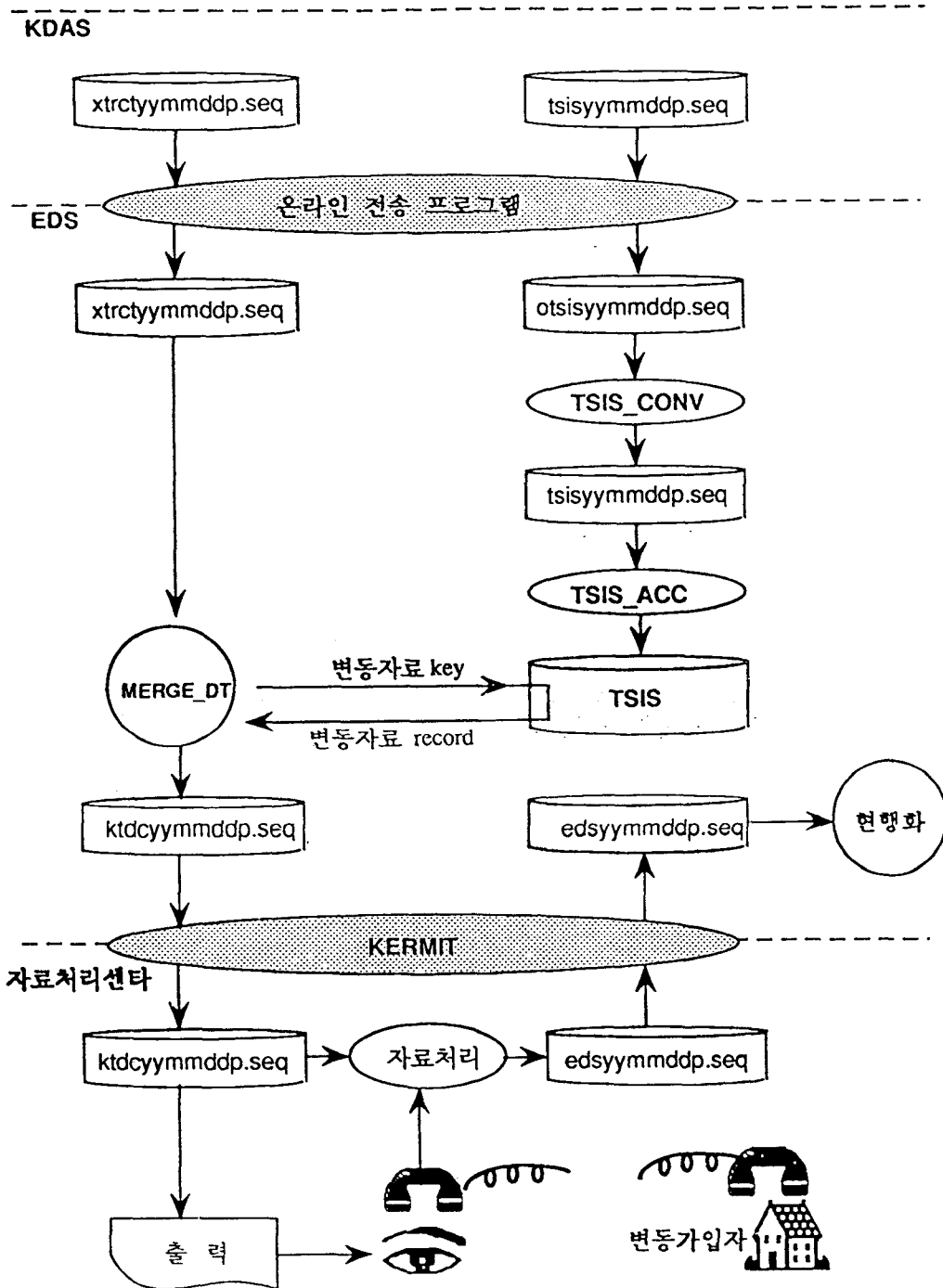
가입자의 변동사항은 영업창구 단말기상으로 입력되어 TSIS에 누적된다. 입력된 TSIS 일일변동자료는 KDAS에 전송되어 안내용의 일일변동자료를 처리 기준일에 따라 추출한다. 추출된 일일변동자료에는 가입자가 원하는 단체명을 안내 성격에 따른 단체명으로 변환함으로 가입자가 원하는 상호명이 상실된다. 그러므로 EDS 일일변동자료에서는 가입자가 원하는 단체명을 수록하기 위하여 TSIS 일일변동자료를 참조해야 한다. 또한 일일변동자료에는 업종코드가 수록되지 않음으로 업종코드가 별도로 수록되어야 한다.

EDS의 일일변동자료는 [그림 3.3]과 같이 KDAS, 자료처리센타와 연동하여 생성된다. KDAS에서 TSIS일일변동자료(otsisymmddp.seq)를 전송받아 KSC5601 2바이트 완성형으로 변환하여(tsisymmddp.seq) CTL화일 관리모듈을 이용, 변동자료키에 의한 색인화된 누적화일(TSIS)에 누적시킨다. KDAS 일일변동자료(xtrctymmddp.seq)를 전송 받아 변동자료키를 검색키로 하여 누적화일에서 TSIS 일일변동자료를 발췌하여 가입자가 원하는 단체명을 추출하여 KDAS일일변동자료에 삽입하여 자료처리센타의 처리전 일일변동자료(ktdcymmddp.seq)를 생성한다. 생성된 일일변동자료는 자료처리센타에서 변동가입자 문의전화에 의한 누락정보및 업종코드를 수집하여 EDS 일일변동자료(edsymmddp.seq)를 생성한다.

4. 현행화

EDS 전화번호 관련정보 화일의 현행화는 매일 발생하는 일일변동자료와 서비스용 화일을 가지고 가입자의 변동정보가 수록된 새로운 서비스용 전화번호관련 정보화일을 유지하는 모듈이다. <표 3.2>의 변동자료의 변동코드를 분석하여 게재구분에 의거 인명, 업종단체 및 친구번호화일을 변경시킨다. 특수화일인 국번과 국내지역화일은 자주 변동이 생기지 않음으로 변동사항이 발생한 경우에만 현행화를 한다.

공유메모리 기법을 사용하여 일일변동자료에 의한 화일을 갱신하면 3~4배의 지연이 소요됨으로 현재는 검색 서비스가 드문 심야(2~6시)에 현행화를 수행한다.



[그림 3.3] EDS 일일변동자료 생성과정

<표 3.2> EDS 일일변동자료 변동코드표

코드	변경내용	세 부 내 용
A1 A2 A4	가입해지 단기가입해지 사업용해지	가입자 자료가 말소되는 경우
B1 B2 B3 B4	일반승낙 긴급승낙 단기가입승낙 사업용승낙	가입자의 전화가 신설 또는 증설되는 경우 <ul style="list-style-type: none"> · 순수신규(해당 계재명이 신설되는 경우) · 신규증설(기존 계재명내에 전화가 증설되는 경우 - 기설전화번호 입력요함)
C1	전입부활	가입자전화가 타 전화국으로 수용구역이 변경되면서 설치장소도 변경되는 경우 <ul style="list-style-type: none"> · 순수전입 · 전입증설(기설전화번호 입력요함)
D1 D2 D3	양도 승계 개명	전화번호에 대한 계재명의 또는 부서명이 변경되는 경우 <ul style="list-style-type: none"> · 다른 단체로 명의변경 · 동일 단체내에서 부서명 변경 · 다른 단체내의 부서로 변경(다른 단체내의 기설전화 번호 입력 요함)
E1 E4 EE	국내설번 지번정정 주소변경	전화기 설치장소가 변경되는 처리
E3	전출	해지와 동일한 처리
I1 I2 K1 K2	대표장치승낙 대표장치폐지 접속장치승낙 접속장치폐지	일반전화를 대표전화로 사용하는 경우 대표전화를 일반전화로 전환 사용하는 경우 일반전화를 팩시밀리로 사용할 경우 팩시밀리용 전화를 일반전화로 사용할 경우
L1 L2 M1 M2 N1 N2	일시철폐 일시부활 통화휴지 휴지부활 통화정지 정지부활	<ul style="list-style-type: none"> · 가입자 자료가 일시 철폐되는 경우 - 해지처리 (L1, M1, N1) · 일시철폐된 전화를 정상적으로 다시 사용하는 경우 - 신규 처리 (L2, M2, N2)
P1 PP	계재명의정정	계재명의를 정정할 경우
Q1	번호변경	가입자의 전화번호가 변경되는 경우
R1	가입원부정정	계재명이 변경될 경우
Q1	종류변경	처리체크값에 따라 구내교환폐지('0'), 신규('1')
T1 T2	수용변경 국번변경	전화국사정으로 수용국, 국번변경이 발생하는 경우

IV. 결론

B-tree구조의 색인화된 순차적 접근방식(CISAM)을 제공하는 CTL화일 관리 모듈을 응용하여 가입자 전화번호 정보화일을 구축, 매일 발생하는 가입자의 변동사항을 수집하여 전화번호정보화일을 현행화하는 EDS 전화번호 관련정보 화일 관리시스템의 개발된 내용에 대해서 살펴보았다.

EDS 전화번호관련 정보화일 관리시스템은 아래와 같은 특성을 갖는다.

- 고정길이 레코드 형태의 색인화된 순차적 접근방식임으로 검색속도가 빠르다.
- 레코드 삭제시 색인키화일의 해당 키값이 삭제가 안됨으로 현행화 수행시간이 빠르다.
- 업종상호인 경우 한 레코드상에서 해당단체의 상하 계층적 조직구조를 알 수 있기에 부서정보의 검색속도가 빠르다.
- 레코드의 한필드상에 복수개 업종코드 수록이 가능하여 효율적인 업종코드 검색 및 관리를 할 수 있다.
- 검색자 위주의 단체명을 보유하고 있다.

EDS 전화번호 화일 관리시스템은 대 용량의 메모리가 요구되지만 신속한 검색 속도, 복수개 업종코드의 효율적인 관리 및 현행화 수행시간이 짧아 공사의 가입자 전화번호 관련정보 화일관리 시스템으로 확장및 효율적인 관리가 가능하다.

[참 고 문 헌]

- [1] KTRC, 전자전화번호부 시스템 개발 89년도 중간보고서, 1989.12.
- [2] KTRC, 전자전화번호부 시스템 개발 90년도 중간보고서, 1990.12.
- [3] KTRC, 전자전화번호부 시스템 개발 91년도 중간보고서, 1991.12.
- [4] KT, 114안내시스템 운용지침서, 1989.10.
- [5] KTRC, 전기통신연구 제6권 2호, 1992.7 .
- [6] CTL, DATABASE LIBRARY FUNCTIONS 메뉴얼, 1990.1.
- [7] MICHAEL J. FOLK, BILL ZOELLICK, FILE STRUCTURES A CONCEPTUAL TOOLKIT, 1987, pp 80 ~ 146.
- [8] 이석호, 데이터베이스론, 1992.2, pp 83 ~ 119.
- [9] 다우, 한글 C-ISAM 참고설명서, 1988.6.