THE TEXAS LOANSTAR PROGRAM: ACQUIRING AND ARCHIVING LOANSTAR DATA

Jeff Haberl, Srinivas Katipamula, Dean Willis, Kristel Weber,
Jennie Matson, Murthy Rayaprolu, Udaya Subramanian
Energy Systems Laboratory, Mechanical Engineering,
Texas A&M University, College Station, Texas.

ABSTRACT

This paper discusses the acquisition and archiving of data for the Texas LoanSTAR monitoring program, an eight year, $98 million revolving loan program for energy conservation retrofits in Texas state, local government and school buildings funded by oil overcharge dollars.

In particular we focus on the design, development and implementation of a state-wide computer network that communicates with field recorders installed in participating agencies, the development of public domain software for electronically polling the data acquisition systems (DASs), and the development of procedures to assess the accuracy of the data being collected. The development of a field data recorder testbench facility is also discussed. Such a facility is being used to develop and test software for polling DASs to be used in the program.

In this report we also discuss data management and the process of acquiring data from a site, translating these data to a common format and preparing the data for graphic presentation or analysis. In the final section we discuss the lessons we have learned and our future work.

OVERVIEW

Data acquisition for the Texas LoanSTAR program is being accomplished by three different means: (1) electronic polling of DASs installed in the participating agencies, (2) manual transcription of data from input forms, and (3) electronic transfer to/from existing data bases. Figure 1 illustrates the three primary data paths. Data to be collected include: site description information, utility billing and rate information, and electronically polled 15-minute and hourly consumption and influencing parameter information (e.g., ambient temperature and humidity, indoor temperature, etc.).

One of the objectives of the LoanSTAR program is to assist building professionals in the state of Texas by improving their access to measured building energy consumption data. Currently, most whole-building or sub-metered building energy consumption data are provided by utility load research recorders that are not always well suited for building energy research. Further, many of the manufacturers of such equipment do not provide their proprietary communications protocols leaving users at the mercy of the manufacturer's software. The net effect is that few building professionals have ready access to monitored building energy consumption data. As a first step toward improving the accessibility of data it was decided that public domain polling procedures and data reduction methods should be developed for several different data recorders, and made available to Texas building professionals.

In order to accomplish this the following sub-tasks were established: (1) the negotiation of licensing agreements with manufacturers in order to obtain the proprietary communications protocols needed to develop the polling software; (2) establishment of a communications testbench for testing the DASs, (3) testing of existing manufacturer's data acquisition system software; (4) developing and documenting procedures for translating data to a common archive format using manufacturer's software; and (5) developing and testing public domain software to poll the DASs.

COMMUNICATIONS TESTBENCH

A communications testbench has been established in the Wisenbaker Engineering Research Center on the main campus of Texas A&M University. Six manufacturers are participating in the program. The characteristics of the thirteen field data recorders from the six manufacturers are listed in Table 1. Three of the manufacturers have signed a software licensing agreement and have delivered their communications protocols.

A schematic diagram of the communications testbench is provided in Figure 2. Three PCs are being used in the testbench; one to generate a pre-programmed signal; a second PC to communicate with the DAS; and a third PC to analyze the communications between the polling computer and the DAS. Currently, the testbench has been set up, the testbench software has been purchased and tested, and the early stages of testing have been initiated with several of the field data recorders.

SOFTWARE DESIGN

Polling the Data Acquisition Systems (DASs) In order to receive the data from the field we have dedicated several PCs to the task of polling the DASs via telephone modem. This is performed weekly, with some interim checks to assure that the systems are still operational between weekly pollings. Each week when the data acquisition systems are called we perform several tasks, including: a check to see if the clock is still on time, resetting the clock if necessary, and the down-loading the data to the LoanSTAR MAP Net (Monitoring and Analysis Program Network). The LoanSTAR MAP Net is a local area network designed to handle the data acquisition, archiving, computation, word processing and graphics needs of the program. It is based on a high performance Unix server and industry-standard Ethernet cabling. The network includes over one gigabyte of disk space, two dozen 386-class workstations, two workstations, two terminals and several postscript quality laser printers. It is directly connected to the TAMU campus-wide network which allows for access to a vast array of computing services.

The Data Recorder Unified Management System (DRUMS) Figure 3 illustrates the conceptual structure of the software for polling, checking and archiving data from the field data recorders -- the Data Recorder Unified Management System (DRUMS). The DRUMS will perform several functions, including: programming of the data acquisition systems, scheduling the polling calls, and translation of commands and data records for different manufacturer's formats.

Common-format data from the DRUMS will then be passed on to the storage processing systems where permanent on-line storage will utilize a Relational Data Base Management System (RDBMS) to facilitate easy retrieval of the heterogeneous data. For each agency participating in the program,it has been necessary to gather: (1) *Point-in-time Information:* such as engineering data, survey information, one-time measurements, interviews, site descriptions, etc.; (2) *Time-Sequenced Information:* including monthly utility billing data, daily minimum-maximum weather data, hourly energy consumption data, etc.; (3) *Influencing Parameter Measurements:* such as ambient temperatures, humidity, solar, wind speed, scheduling information, etc.; and (4) *System Requirements Information:* such as design information, environmental quality requirements, comfort requirements, lighting conditions, etc.

Typical Data Path for LoanSTAR Agency In order to rapidly facilitate the development of software for the project, we decided to use a modular approach, and as much as possible, "canned" graphic and statistical programs that allow us to string together numerous small modules, or "filters" to accomplish a larger task. Figure 4 is an example of our programming flowcharts. In the Bibliography we have listed the software packages that we are using in the program. In general they consist of specialty columnar-data manipulation modules from academic and freeware sources, and various proprietary graphics and statistical analysis packages.

Briefly, here are some of the steps that are performed in the flowchart in Figure 4. After the site is polled the first step is to remove extraneous characters from the manufacturer's ASCII data. For example, the data from one manufacturer contains "/", ":", and "A-Z" characters as well as numbers in "-123E-10" format. In order to process this with the ARCHIVE software (from Princeton University) we have had to do some preliminary pre-processing to remove these extraneous characters from each hourly record. This is performed by the RAW2DAT.AWK routine, a PC-awk script. Awk is a Unix-based programming language developed by the AT&T Bell Laboratories.

Next, we pass the data to the ARCHIVE program (with the appropriate channel table for each site) and produce log reports and columnar ASCII data that is almost ready for archiving on the Unix server. One additional step is needed whereby the MISSING.AWK program inserts a "missing" character (i.e., -99) for whole records of data that are not contained in the original files. Data are now ready for loading onto the Unix hard disk (via the Ethernet) or can continue on for additional processing.

The remainder of the flowchart contains four additional processing paths that produce derived weather channels and pre-process the data for plotting. The path heading directly down from the .ACS file through the RHTRIM.PAS procedure produces relative humidity values that are trimmed at 99.9999 whenever the incoming data registers 100.5 -- an indication of a saturated humidity sensor.

The path that passes down through AIR.PAS routine calculates enthalpy and absolute humidity given dry-bulb temperature and relative humidity. This path is then recombined with the original weather channels and stored on disk for each weather site in *.WEA file.

The paths that pass through PLUCK.AWK and through AHU.DAT generate plotting information for the weekly inspection plots. Figure 5 is an example of the weekly data inspection plots that are produced to view each of the respective channels. For the most part each of the graphs represents one channel of data or one channel plotted against another. Figure 6 is an example of a derived graph which is viewed on a regular basis. This graph presents one month of hourly whole-building, sub-metered and derived electricity information from 12 channels of electricity data. The upper line represents the whole-building electricity data recorded at the main service panel in the building. The second line from the top is a derived channel which is representative of all other electricity consumption in the building, for the most part lights and electrical receptacle loads. The next line represents the electricity used by the motor control centers (primarily motors in the air-handling units). The lowest line represents the electrical energy use of the large computer center in the Zachry building.

Availability of LoanSTAR Software Our two primary objectives with developing the LoanSTAR software are to: (1) collect, analyze, archive and distribute energy consumption data for those agencies participating in the program; and (2) develop and disseminate public domain building energy analysis tools that perform such processing. With this in mind we have made provisions for licensing and distributing software through the Texas Engineering Experiment Station (TEES) Office of Contract and Grants. Software produced for the LoanSTAR program will be public domain and will be available for a modest distribution fee.

The following modules have been produced as part of our work during this first year and are being prepared for distribution: (1) A columnar-to-matrix pre-processor for 3-D plots which utilize the Intex Solutions 123 add-on package. (2) Automatic macro routines for processing the 3-D plots with the Lotus 123 program. (3) An hourly ASCII data extraction program for DOE-2. (4) A file handling batch routine for PRISM runs that utilizes multiple weather stations. For more information on PRISM see Fels et al. (1986). (5) Psychrometric routines for calculating absolute humidity and enthalpy given dry bulb temperature & relative humidity. (6) A "missing data" add-on awk script for ARCHIVE. (7) A "cleaning" awk script pre-processor for ARCHIVE that removes "/", ":", etc., from Synergistics C180 data. (8) Procedures for preparing Sangamo data for ARCHIVE.

LESSONS LEARNED

The first year's experience of undertaking a large-scale energy monitoring project has taught us many lessons. Here are some of the more important issues involved with acquiring data: (1) Many data acquisition manufacturers are hesitant to release their communications protocols. Without such protocols, analysts are locked-in to proprietary software and data formats that may not be suitable for their analysis. (2) There is a need for affordable, public domain data analysis toolkits that are powerful, flexible and can be applied to any data stream. (3) A large (i.e., multi-tasking, or Unix-based) data storage system is needed to monitor hundreds of hourly channels and routinely prepare reports for dozens of buildings. DOS-based systems suffice, and are easier to implement, but quickly run out of memory and disk space, and end up costing about the same as a Unix system.

FUTURE WORK

The future work during the 1990/91 year will evolve around securing additional Software Licensing Agreements, determining the feasibility of acquiring hourly weather data over the Internet from the National Weather Service, continuing with the development of the testbench and DRUMS, developing the relational data base for the Unix server, exploring the feasibility of monitoring building energy usage with Energy Management and Control Systems, and taking a closer look at using portable monitoring systems.

List of Data Acquisition Systems for Bench Test

Manufacturer	Unit	kW	Digital	Analog
Synergistics Control Inc.	Datamate 10● (Level 1)	-	Four	-
	Datamate 20● (Level 1)	-	Eight	-
	Datamate 30● (Level 2)	-	-	Eight (4-20ma) (0-5vdc)
	Datamate 40● (Level 2)	-	Four	Four (4-20ma) (0-5vdc)
	Datamate 50● (Level 2)	-	Eight	Eight (4-20ma) (0-5vdc)
	C-140● (Level 2)	Six	Six	Six (4-20ma) (0-5vdc)
	C-180 (Level 3)	Sixteen	Sixteen	Fifteen (4-20ma) (0-5vdc)
Campbell	21X (Level 2/3)	-	Four	Eight/Sixteen Options*
	CR10● (Level 1/2)	-	One/Two	Six/Twelve Options*
Rustrak	Ranger II (Level 1)	-	Four	Four
Landis & Gyr	DataGyr 100 (Level 1)	Four	Four (2/3 wire)	-
Process System, Inc.	Sentry 200 (Level 1/2)	Four	Four 3 wire Eight 2 wire	-
Sangamo Systems	ST-DS111 (Level 1)	-	One/Two/Four	-

●Proposed
*Doesn't include 4-20ma

Table 1: LoanSTAR DAS Vendor Status. *This table lists the data acquisition systems that are under evaluation for the LoanSTAR program. The unit label refers to the manufacturer's model; kW refers to the number of direct current transducer (CT), potential transducer (PT) channels that the unit will accept; the Digital column designates how many digital inputs the unit will accept; the Analog column refers to the number of analog input channels the unit will accept.*

BIBLIOGRAPHY

Fels, M. (ed.) 1986. "Special Issue Devoted to Measuring Energy Savings, The Princeton Scorekeeping Method (PRISM)", *Energy and Buildings,* Vol. 9, Nos. 1 and 2.

Free Software Foundation, 1989. *GAWK*, (the PC version of the Unix-based AWK toolkit), 675 Massachusetts Ave., Cambridge, MA, 02139.

Golden Software, 1990. *Grapher & Surfer*, 809 14th Street, P.O. Box 281, Golden, Colorado, 80402-0281.

Intex Solutions, 1990. *3D Graphics,* 161 Highland Ave., Needham, MA, 02194 (requires Lotus 123).

Lantern Corporation, 1990. *Voyager: Data Exploration Software,* 63 Ridgemont Drive, Clayton, MO, 63105 (requires Microsoft Windows).

Princeton University, 1987. *ARCHIVE: Software for Management of Field Data,* D. Feuermann and Willett Kempton, Center for Energy and Environmental Studies.

Princeton University, 1987. *Tony's Tools,* Anthony Lovell, currently at Yale University, (comes with the ARCHIVE program), Center for Energy and Environmental Studies.

Princeton University, 1987. *Art's Tools*, Art McGarrity, currently at Swarthmore College, (comes with the ARCHIVE program), Center for Energy and Environmental Studies.

SAS Institute, 1990. *PC SAS,* SAS Circle, Box 8000, Cary, North Carolina, 27512-8000.

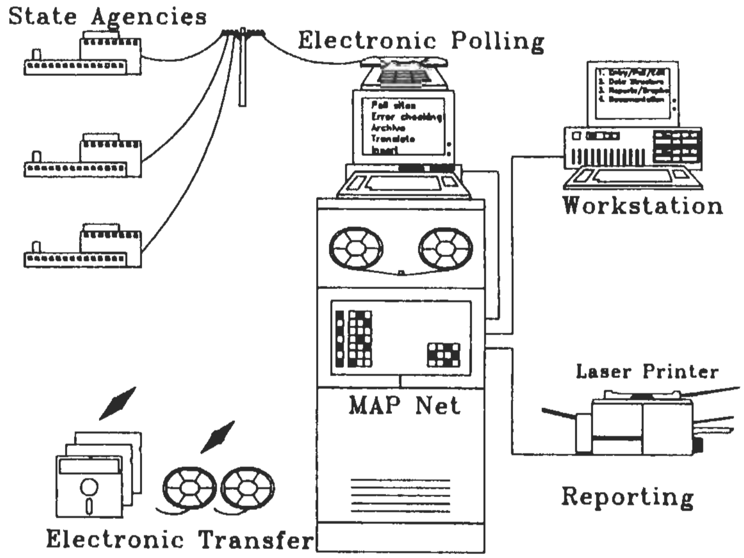

Figure 1: Data Paths for the LoanSTAR Program. The three primary data paths for LoanSTAR Agencies are shown in this figure. Data can be polled electronically, manually entered, or transferred to/from external databases.

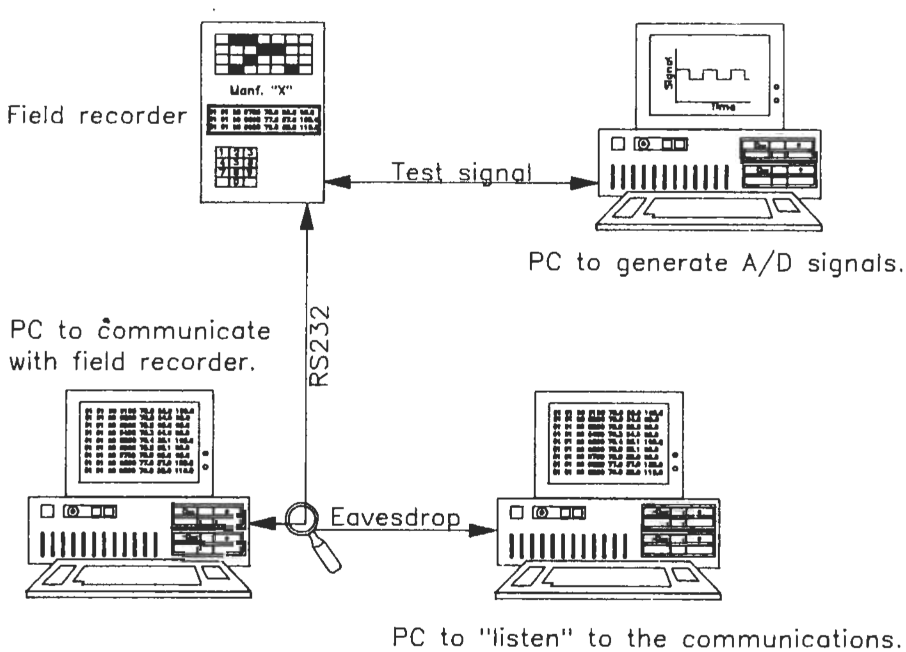

Figure 2: The Communications Testbench. This figure illustrates the communications testbench that has been established. The testbench is used to verify the accuracy of the data acquisition systems (DASs) and to develop the polling software. Three PCs are being used in the testbench; one to generate a pre-programmed signal; a second PC to communicate with the DAS; and a third PC to analyze the communications.

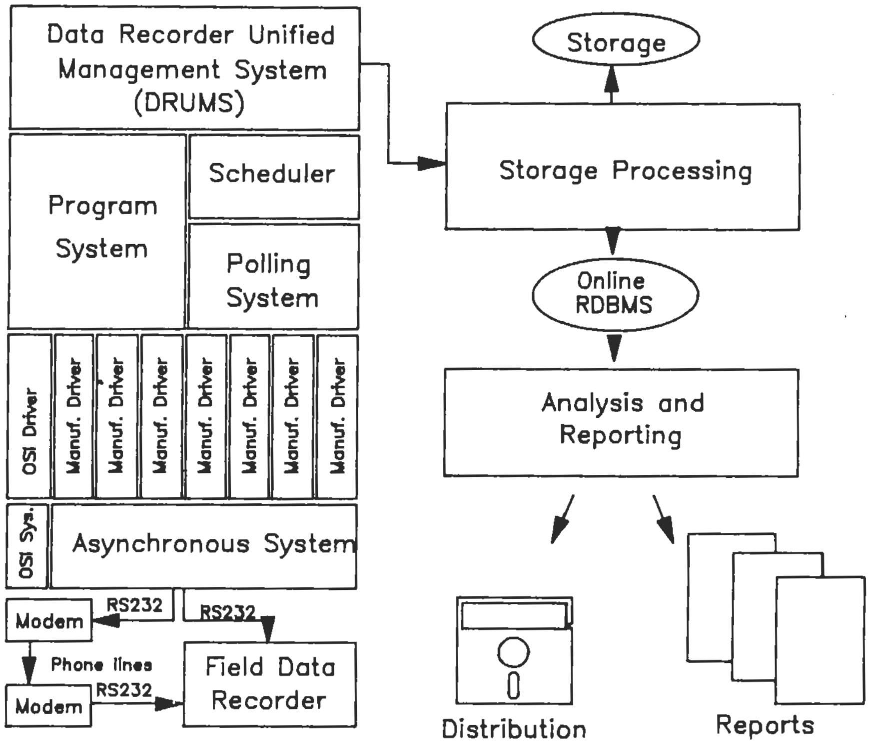

Figure 3: The Data Recorder Unified Management System (DRUMS). A schematic figure of the DRUMS is shown here. Data are retrieved periodically from various sites via modem using the appropriate manufacturer's driver. Once the data are translated to a common format, they will then be stored for analysis and reporting in a Relational Data Base Management System (RDBMS).

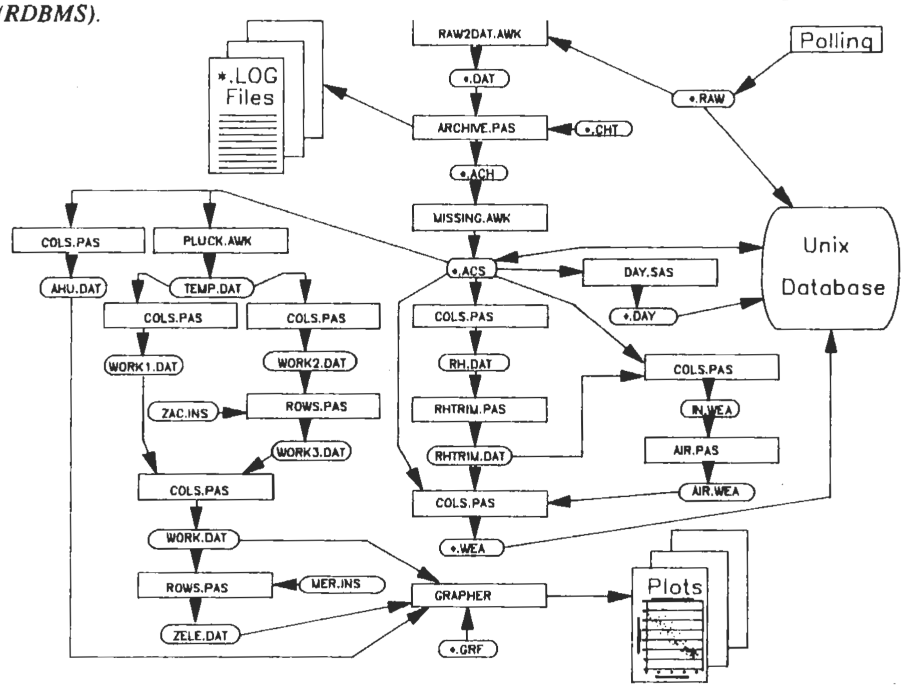

Figure 4: Data Path for the Zachry Engineering Center (Z.E.C.). This figure is an example of the data path for the Z.E.C. Data are polled once per week and archived in both raw and processed format. Various modules have been developed for removing extraneous characters, checking for missing data, calculating derived weather channels, and producing various graphs.

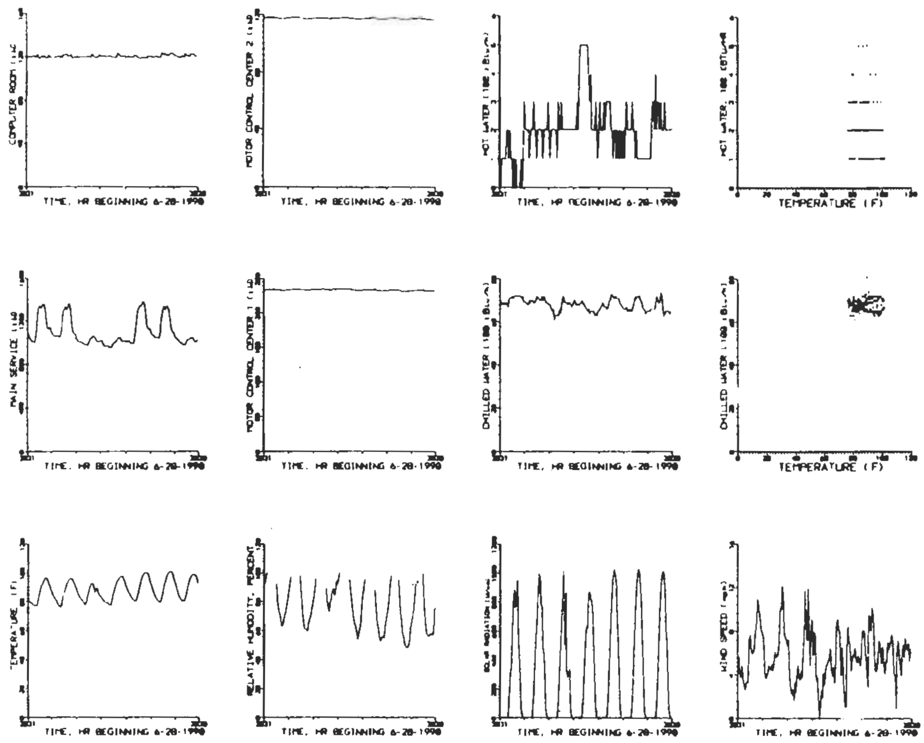

Figure 5: Weekly Verification Plots for Zachry Engineering Center. *This figure is an example of the weekly verifications plots for the Z.E.C. Each graph represents a channel of information plotted in time-series or versus another channel of information. These graphs are used for visually inspecting the data that are coming from the sites.*

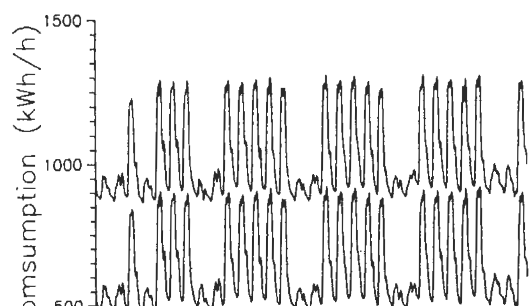

Figure 6: Derived Electricity Plot for Zachry Engineering Center. *This figure displays whole-building, sub-metered and derived electricity consumption for the Z.E.C. Starting from the top, the upper line represents the whole-building electricity use, the next line represents derived energy use for the lights and electrical receptacles, the third line from the top represents energy used by the motor control center and the lowest line represents energy used by the computing center.*