EXPERIENCES WITH FAX DATA REPORTING AND QUESTIONNAIRE DISTRIBUTION

Martin V. Appel, Thomas F. Petkunas, Chad E. Russell, U.S. Bureau of the Census
Chad E. Russell, U.S. Bureau of the Census, Washington, D.C. 20233-0001

Introduction:

The Census Bureau has long used traditional mail-out/mail-back methods to conduct surveys. Questionnaires are mailed to respondents, who complete them and mail them back. The data is then keyed into a computer file for further processing. The Paperless Fax Image Reporting System (PFIRS) was conceived as a system whereby respondents could return their completed questionnaires via facsimile. Data on the fax images would then be automatically interpreted through intelligent character recognition (ICR) and converted to ASCII output.

To test this technology, the Census Bureau assembled a proof of concept system with commercially available hardware and software products and began testing on fifty Survey of Manufacturer's Shipments and Inventories (M3) respondents. Each month, these respondents would fax their completed questionnaire to an 800 number. The data was extracted using ICR software and the accuracy of the results observed.

Other applications of the system's fax capability were also explored. For the Survey of Industrial Research and Development (R&D), and the Investment Plans Survey (IPS), replacement questionnaires could be requested via a Touchtone Data Entry (TDE) system. Respondents needing a questionnaire called an 800 number and placed a request using their telephone keypad. The PFIRS system would then build and send a questionnaire by fax to the number they specified. In addition, we also tested using the fax out capabilities to send reminder notices to delinquent respondents in the IPS survey.

This paper describes our experiences with the proof of concept system during a year of testing. We will examine the accuracy of the recognition process, the ease of use, and the success of the current system with respect to outgoing fax applications.

Description of the System:

The proof of concept system resides on a small IPX/SPX network consisting of a fileserver, a dedicated fax server containing eight fax/modem boards, and a 486 PC running the ICR/forms management software (Teleform 3.0). Teleform is a comprehensive software package that includes modules for forms design, ICR, fax management, and verification of the extracted data. The eight fax/modems are connected to eight analog telephone lines in a hunt group connected to a toll free number.

In order for the ICR capabilities of this particular software to function, the questionnaire to be interpreted must be created using the form design module. Once the form has been defined, it can be printed or faxed to the respondent. The respondent then completes the questionnaire and faxes it back to the toll free number.

Once the fax image is received by the fax server, it is automatically forwarded to the ICR workstation, where the image is deskewed and identified. If the image quality is sufficient for the software to identify the form, the data is then interpreted by the software's ICR engine. Teleform is bundled with the Nestor ICR engine, one of the more popular hand print recognition engines commercially available. The software does not attempt ICR if the form cannot be identified.

In order to ensure accuracy, confidence thresholds for ICR interpretation can be defined for each data field at the form design stage. If the recognition engine cannot interpret a character with sufficient "confidence," the field is flagged for human verification. By setting these confidence levels, one can control, to some extent, the accuracy of the results of the data capture process.

The verification module provides a graphical user interface for checking/correcting the results of ICR, allowing the verifier to key unrecognized characters by hand. Once the data has been verified, it can be exported in a number of database or spreadsheet formats for further processing.

With respect to sending faxes, the software has the capability of overlaying data from external database files onto forms for printing and faxing (called a form merge). This allows forms to be "customized" for the recipient with company name, census ID, etc. The information to be merged must be arranged in a database file having certain key fields. This database is then associated in a one-to-one manner with an outgoing form. The printing or faxing of the merged forms can be done on a flow basis, or can be scheduled through a time stamp variable.

Test Description:

The proof-of-concept was designed with the

following questions in mind: 1) how accurately will the system recognize the data on a fax image; 2) how much storage space will be required to hold the image files; 3) will Census forms need to be modified in order to obtain adequate recognition levels; and 4) how many phone lines are necessary to conduct a survey of a given size?

The Survey of Manufacturer's Shipments and Inventories (M3) was chosen as a pilot test. The M3 is a short (depending on the industry, either five or seven data items) monthly survey for which respondents report various aspects of orders and inventories in dollar amounts. The M3 staff has a history of innovation, and was already receiving approximately one third of all responses by fax.

For the test, 100 establishments who consistently report by fax were hand picked (not a random sample) to participate. Each month, the test group was mailed a PFIRS version of the standard M3 questionnaire and asked to fax their response to the PFIRS 800 number. Upon receipt, each form was 100 percent verified to evaluate the accuracy of ICR. Statistics collected from receipts included number of characters detected, number correctly interpreted, whether interpretation was above/below the confidence threshold, number of forms not identified, the presence of line noise or interference, date of receipt, image size, and others.

Form Identification:

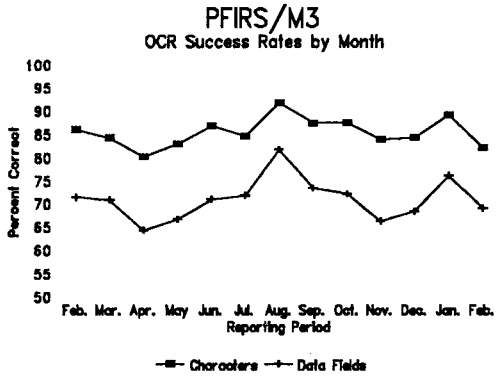
For the February 94 through July 94 reporting periods, of the 214 questionnaires received, the software did not recognize 5 (2.3%) of the images received as questionnaires. For August 94 through February 95, of 523 forms received, 16 (3.0 %) were not recognized. (NOTE: Beginning with the August 94 reporting period, the software was upgraded to a new version containing an update to the recognition engine. Since the recognition engines being applied are different, statistics are reported separately).

Some success has been noted with reevaluation; that is, it is sometimes possible to get a successful ICR pass by sending the nonform through the reader a second time. One possible reason for this is that repeated applications of the deskewing transformation can result in eventual form identification, but this has not been verified.

Accuracy of the ICR Interpretation:

Figure 1 shows the unconditional success rate of ICR on both a character basis and a field basis. This is the percentage of characters/fields interpreted correctly by the software, regardless of the level of confidence at which the characters or fields were interpreted. One can think of this as the proportion of the characters or fields that would be correct if no verification of the data was performed whatsoever.

Figure 1



For the first six months of testing, the average unconditional success rate, defined as characters interpreted correctly by the software, was 84 percent. For the seven months following the software upgrade, the average ICR success rate was 86.5 percent. As one would expect, the slightly better recognition rate is reflected in the field success rate - 69.1 percent versus 72.4 percent.

Table 1 shows the estimated conditional probability that a character is flagged as questionable, given that it is interpreted incorrectly. Different amounts of data were collected at the various confidence levels, thus the level of precision in these averages is different. Generally, as the confidence level is lowered, the probability that the interpretation is below the confidence level is lower (we consider the first entry in the second column to an outlier).

Table 1: P (Flagged|False)

Confidence Level	Version 2.0	Version 3.0
Max	64.2	65.6
95	49.3	76.4
90	34.2	69.0
85	N/A	45.1

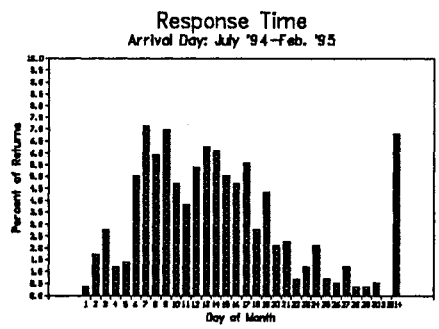
The improvement of version 3.0 over version 2.0 is also seen in Table 1.

Response Rate and Response Time

The overall average response rate for the M3 test group was 82.5 percent. This figure includes late returns - those coming in after the subsequent month's mailout, and is adjusted to account for different test sample sizes in each month.

When we started tracking the arrival time of late arrivals, approximately 93 percent of the forms were returned before the next month's mailout. The remainder would continue to trickle in during the following months. Figure 2 shows the frequency distribution of the return date for the July '94 - February '95 late arrivals. Eighty-five percent of returns are received by the end of the third week of the month.

Figure 2



Storage Requirements:

The required storage space for each image file improved greatly with the release of the software update, which made use of TIF compression. Prior to the update, PCX format was used. Average image size went from 300 Kb to 50 Kb per questionnaire.

Ease of Use:

Correction of returned questionnaires compared favorably with keying in preliminary time tests. On a test set of 80 questionnaires, a verifier was able to do 100 percent verification in 26 minutes, compared with 23 minutes for a seasoned data entry person. The graphical user interface is cumbersome for some tasks. The real savings comes in when you don't have to correct fields. Data entry costs (in keystrokes) decrease linearly as ICR improves.

Further Testing:

More "production-like" testing began in May 1995. Rather than performing 100 percent verification on the data items, the software was allowed to skip verification of items which it had interpreted with sufficient confidence (95 percent confidence level). Received faxes were also keyed by an experienced keyer. The keyed data was matched against the data that had gone through ICR and verification, to locate the differences. Before correcting the errors in the data, it was allowed to pass through the routine automated edits that are applied to all M3 data. The goal was to determine whether the errors that remained after ICR and verification would be detected by the automated edits, and to determine the final "output" error from ICR/verification.

Thus far, three months of data have been collected. Of the 1035 data items processed, only 10 incorrect items passed through the ICR/verification process and were not flagged by the M3 edit program: an output error rate of approximately 1 percent.

Results from the R&D and IPS Applications

The R&D and IPS surveys provided an opportunity to test two outbound fax applications: sending reminder notices and fax on demand of replacement questionnaires. For the fax on demand applications, persons needing a replacement questionnaire could call toll free and enter their ID and fax number using their telephone keypad. The TDE system would then write this information to a form merge database, and PFIRS would automatically build and send the replacement questionnaire.

IPS Reminder Notices:

The goal for the reminder notice application was to fax a reminder notice to delinquent respondents for whom we could obtain a valid fax number. The workload for the 1994 survey cycle was approximately 10,000 cases, to be completed in one week. Results were disappointing. Due to various

problems, only 41 hours of uptime were achieved, during which 2651 reminder notices were faxed (an average throughput of 59 reminder notices per hour). The average transmission time for each was 2 minutes.

One of the lessons learned from this is that final throughput is a function not only of image production time, fax transmission time, and number of outgoing telephone lines, but also time that the software spends on "administrative" tasks. Despite efforts to reduce this time, these lags prevented all eight outgoing lines from being used efficiently. The software was probably not designed with extremely high volume in mind.

Throughput was greatly improved in the second survey cycle by running the software on two machines simultaneously. During the second survey cycle, 6991 reminder notices were faxed in seven days. By using two machines, images were created fast enough to accumulate in the outgoing fax queue, and all eight lines were used consistently. Throughput was approximately 145 reminder notices per hour.

Replacement Questionnaires -R&D and IPS

For both the IPS and the R&D survey, the PFIRS system was integrated with the TDE system to provide fax on demand of replacement questionnaires. For the IPS survey, approximately 13,000 people were sent reminder notices (faxed or mailed) informing them of the availability of this service. Between December 8, 1994 (when the reminder notices were mailed) and the beginning of January, 1,794 requests were received; of these, 1,587 questionnaires were successfully sent. By the close of the operation, there had been 2,018 successful sends. The requests not successfully sent include cases where the fax line was consistently busy, or the respondent entered a voice telephone number by mistake.

For the R&D survey, the replacement form service was available for both the 1994 and 1995 survey cycles. The total 1994 workload for the year (from June, 1994 through January, 1995) was approximately 4,000 requests. During the first two weeks of the R&D operation in 1994, more than 1,200 requests were received, causing the outgoing fax queue to fill up with images. This problem was corrected for the 1995 survey cycle. The initial demand after the first 1995 reminder notice mailout was much less - only 643 sends in April 1995.

One reason for the apparently higher 1994 demand was the TDE voice prompt. Callers were told that they would receive their fax in "about an hour." Because the system was operating on a "last in, first out" basis with incoming requests, it is likely that many callers called more than once after not receiving their fax within the hour. The voice prompt was changed for subsequent survey cycles, promising receipt within "twenty four hours." This allowed the system, even if saturated with requests, to clear the queue of requests during the nighttime hours.

Feedback from the R&D and IPS staff has been very positive. Despite initial problems with the R&D survey, the replacement form applications worked very well, even though the software was not designed with fax on demand in mind.

Conclusions:

Based on these experiences with the proof of concept system, the accuracy of ICR has reached the point where substantial savings of keying costs can be realized. These tests have shown where the current system is lacking (image indexing, insufficient throughput, necessity of recreating a new form version, etc.) so that we can better understand what is needed in a production prototype.

Originally this project was conceived for data collection, so the demand for outgoing fax applications was very surprising. The cost of preparing mailing packages made the fax on demand application very appealing. According to survey staff, the first year of the R&D application saved an estimated 88 person days. Outgoing fax applications will clearly have an important place in the production system.

During the development of a production prototype, the proof of concept system will continue to be used for small to medium sized applications. While it is still in use, we continue to look for ways to increase its throughput for outgoing fax applications, and experiment with alternative means of printing documents for ICR.

These experiences have presented a much clearer picture of how a production system of this type should function. As a next step, hardware/software products that will provide these basic functions will be examined, and the weaknesses of the proof of concept system will be improved. Approximately fifteen companies responded to requests for information (RFI) on such products. Vendors of established systems, such as the TDE and the new data keying system, may have products that provide some of the functions of PFIRS. Such products, if functionally adequate, would likely be easily integrable with these existing systems.

References

Rowe, Errol and Appel, Martin. "Paperless FAX Image Reporting System (PFIRS)." ASA, 1993.

Rowe, Errol and Appel, Martin. "Image Processing of Facsimile Data Reporting, Initial Technical Assessment." December 2, 1992.

Russell, Chad. "PFIRS Six Month Report." September, 1994

Russell, Chad and Appel, Martin. "Experiences with the Paperless Fax Image Reporting System." May, 1995