CASE HISTORY:
Effects of Operator Corruption on System Data Integrity

Bradley J Brown

ABSTRACT

Integrity is frequently a highly critical concern for computing systems, especially those which deal with financial concerns and in areas with high legal liability exposure in the event of loss or corruption of data. Systems for which data integrity is a high concern must consider system operators as a possible source of data corruption. A Case History is presented which was developed as a result of a personal interview with the victims of an occurrence of operator corruption. It gives the history of that occurrence and suggests a possible approach to reducing the identified avenue of data corruption.

CASE HISTORY

During the 1990 Income Tax season, a high volume tax preparation company--referred to herein as RWTax--experienced several incidents of operator-induced data corruption to their Electronic Filing System data. The Electronic Filing System is a recently introduced method by which income tax practitioners can transmit tax returns electronically to the Internal Revenue Service. As is to be expected when dealing with taxes, the tax practitioner is responsible for maintain controls in place to eliminate the occurrence of data corruption and is subject to a variety of sanctions by the IRS in the event data corruption occurs.

RWTax had decided to offer Electronic Filing services to their 4500+ customers in sufficient time to allow graceful introduction of the service including evaluation of several Electronic Filing Network companies and training of their personnel. Unfortunately, the company selected for implementation was unable to handle the volume of business encountered and, after losing nineteen returns somewhere between their computer system and the IRS, decided to file for bankruptcy. In the scramble to locate another Electronic Filing Network provider, RWTax discovered that the best alternate vendor's system could not directly access the data from RWTax's existing tax preparation computer system. Given the urgency of the moment, the most appropriate solution was felt to be that of hiring a system operator to handle the new Electronic Filing Network companies' system including the manual transcription of data from RWTax's computer system to that of the Electronic Filing Companies.

During the subsequent weeks the owners of RWTax discovered that the system operator, who had prior experience as a data entry operator, was routinely entering fallacious data into the

system in order to maintain productivity and reduce personal inconvenience.

Social Security Numbers for Children. One instance of operator corruption of data occurred whenever the Tax Return specified dependant children over the age of two that had not received a Social Security Number. Rather than take the time necessary to retrieve the required information, the system operator would merely indicate that the child was under the age of two, thus eliminating the requirement for entry of a Social Security Number.

Partial Custody of Dependant Children. Another instance was that of dependant children who had lived with the Tax Payer for less than six months during the year. Rather than retrieving and/or preparing the required supplemental forms, the system operator would merely indicate that the dependant child was subject to a pre-1985 divorce decree which is a provision not requiring additional forms. This entry was made even for taxpayers who had not even received their divorce yet by 1985.

Reported Income Discrepancies. The third instance was that of forced acceptance of reported income discrepancies. Since the IRS required the presentation of all tax return monetary data to be rounded to dollars, whereas clients who prepared their own returns and brought them in to be filed without further computation on the part of the tax practitioner frequently calculated the returns to the penny, the Electronic Filing Network software would identify the existence of a monetary discrepancy between the reported W-2 wages and the income amount entered on the Income Tax Form 1040, then provided for the forced acceptance of that discrepancy such as would be desired if the discrepancy resulted from rounding to dollars. The system operator however, apparently mis-understood the purpose of the Forced Acceptance over-ride and exercised that function whenever the income discrepancy exception encountered. A casual encounter revealed the case where a income discrepancy of $16,000 resulting from the tax practitioner overlooking one of many W-2 income reporting forms, could have been transmitted to the IRS as a result of the Forced Acceptance over-ride (although it wasn't almost purely by chance).

These three instances of system operator-induced data corruption could have resulted in significant financial impact for the RWTax Company. The worst case scenario resulting from IRS detection of these instances of data corruption would have been substantial monetary sanctions as well as debarment from providing Electronic Filing Services. Debarment from providing services would have represented a significant loss of business and possibly client good-will even if the debarment were only for a limited duration.

REDUCTION OF RISK

Although it is not held possible to completely prevent similar instances of system operator-induced corruption, it is held that Multiple Source Verification of critical data entries could reduce the occurrence of data corruption. In every instance of critical data entry, a secondary entry of data which is isolated temporally from that of the primary entry, should be utilized to verify the accuracy of the primary data input.

In the instance of Social Security Numbers for Children, the birthdate of each dependant child should have been entered as part of the initial tax form data entry, preferably entered into a master client record which is maintained separately from the tax form data itself. When a dependant child does not have a Social Security Number, the system could then verify the age of the child itself from the master data record.

In the instance of the Partial Custody of Dependent Children, the master data record for the client would contain the date of any divorce decree. In the event the pre-1985 divorce decree option is selected, the master data record would be examined to determine whether the divorce date was in fact pre-1985.

In the instance of Reported Income Discrepancies, the amount of allowable forced income discrepancy would be limited to less than one dollar. Any discrepant amounts greater than one dollar would require the approval of the tax practitioner for continued processing.

CONCLUSIONS

Personal interviews with the owners of the company referred to herein as RWTax revealed that their speculation as to the cause of the operator's actions is that of the operator's basic personality. The Electronic Filing system operator had been hired as an 'expert' with computers, was extremely eager to maintain that image. This desire was apparently combined with a basic personality type that found admission of error unpleasant. This situation was possibly exacerbated by the aggressive and forthright personality of one of the owners, which would have served to make it even more difficult on the part of the system operator to ask questions and admit error.

An additional consideration was the operator's apparent misconception that if the computer system accepted a particular data entry, that the system had checked the input for errors and found the data acceptable. Although this was in fact true for syntactic errors, the operator erroneously adduced that the system checked for all valid combinations of data. The operator felt that if the system would accept the absence of a Social Security Number by entering an indication that the child was under two years old, that the system had somehow checked that information

and found it correct. Apparently the operator's conceptual model of the computer held that the computer could somehow accumulate information independently of the operator's own data entry for use in determining validity of data. Although this conceptual model is obviously incorrect to any one even a modicum of knowledge about computer operation, this operator had apparently received all of his prior computer-related training from a temporary service whose training personnel were themselves relatively computer illiterate.

Operator corruption of data entry has been reported before by Cook1 in her recital of a data entry operator who deliberately omitted telephone numbers in an effort to increase personal productivity. It is postulated that the occurrence of operator induced data corruption is much more wide spread than has been reported in the literature. This case study is not intended as an indictment of any individual or occupational group, but rather to alert the Computing Security community of a potentially extensive threat to data integrity with corresponding effects on system security.

-- NOTES --

1. Cook, Janet M. "What C.S. Graduates Don't Learn About Security Concepts and Ethical Standards: Or - 'Every Company has Its Share of Damn Fools. Now Every Damn Fool has a Computer.'" SIGCSE 18.1 (1986): 89-92.

Bradley J Brown
3945 Sonora Ct., Columbia, MO 65201, (314) 442-5881