

MỘT SỐ VẤN ĐỀ RÚT RA QUA XỬ LÝ SỐ LIỆU TỔNG ĐIỀU TRA CÁC CƠ SỞ KINH TẾ, HÀNH CHÍNH, SỰ NGHIỆP 1/7/2002

Phạm Thế Năng^(*)

Xử lý số liệu tổng điều tra là một công đoạn quan trọng trong qui trình tiến hành một cuộc tổng điều tra. Nếu thực hiện công đoạn này tốt sẽ góp phần nâng cao chất lượng, hiệu quả, rút ngắn thời gian công bố và tạo điều kiện thuận lợi cho người sử dụng tiếp cận với kết quả tổng điều tra.

Trong hơn 10 năm qua, Tổng cục Thống kê đã tiến hành 2 cuộc tổng điều tra cơ sở kinh tế hành chính sự nghiệp, theo đó là 2 lần tiến hành xử lý số liệu cuộc tổng điều tra này.

1. Khối lượng thông tin đã xử lý trong tổng điều tra cơ sở KTHCSN 2002

Phiếu số 1: 56.578 phiếu điều tra - 33 trang số liệu

Phiếu số 2: 44.014 phiếu điều tra - 3 trang số liệu

Phiếu số 3: 152.055 phiếu điều tra - 4 trang số liệu

Phiếu số 4: 2.619.341 phiếu điều tra - 2 trang số liệu

Tổng số trang số liệu cần xử lý: 7.846.018 trang số liệu.

Tổng dung lượng thông tin lưu trữ số liệu gốc: 1050 MB

2. Mô hình xử lý

Nhập tin phân tán tại các địa phương và xử lý tập trung tại Trung tâm Tin học Thống kê (TTTHTK):

- 64 Cục Thống kê tham gia nhập tin các phiếu điều tra của địa phương theo chương trình phần mềm thống nhất do TTTHTK xây dựng, chuyển dữ liệu nhập tin về TTTHTK qua đường truyền tin CUTFTP. Thời gian từ tháng 11/2002 đến tháng 3/2003.

- TTTHTK nghiệm thu số liệu nhập tin các địa phương, xử lý, tổng hợp kết quả cho từng địa phương và cả nước từ tháng 4/2003 đến 7/2003.

- TTTHTK chuyển giao dữ liệu sau khi đã làm sạch, chương trình xử lý tổng hợp và kết quả ghi trên đĩa CD ROM tới từng địa phương tháng 8/2003.

3. Tổ chức tập huấn nhập tin

TTTHTK đã tổ chức 3 lớp tập huấn nhập tin vào tháng 10/2002 tại:

- Hà Nội: cho các tỉnh, thành phố phía Bắc
- Đà Nẵng: cho các tỉnh, thành phố miền Trung, Tây Nguyên
- Tp. Hồ Chí Minh: cho các tỉnh, thành phố phía Nam.

Các lớp tập huấn tại 3 miền đã giải quyết một số vấn đề sau:

- Thống nhất qui trình xử lý chung áp dụng cho cả nước.

- Giải đáp các vướng mắc của các địa phương trong quá trình xử lý phiếu điều tra trên máy.

^(*) Trưởng phòng Trung tâm Tin học Thống kê

- Tiếp thu các ý kiến đóng góp của các địa phương về nghiệp vụ liên quan đến quá trình xử lý.

- Hoàn thiện chương trình nhập tin, kiểm tra thông tin và chuyển giao chương trình chính thức tới 64 cục thống kê qua đường truyền tin để các địa phương tổ chức nhập tin.

4. Kiểm tra, giám sát quá trình nhập tin

TTTHTK phối hợp với tổ thường trực của Ban chỉ đạo tổng điều tra của Tổng cục, kiểm tra, giám sát một số tỉnh, thành phố phía Bắc tham gia nhập tin, đặc biệt là Hà Nội. Qua việc kiểm tra giám sát này đã phát hiện một số vấn đề còn tồn tại trong quá trình thu thập, kiểm tra, ký mã của các địa phương như:

- Khái niệm đơn vị điều tra chưa thống nhất cho các tỉnh, thành phố:

- Một số các chỉ tiêu để phân loại cơ sở chưa thống nhất cho các tỉnh, thành phố.

- Mã ngành hoạt động của cơ sở chưa được thống nhất trong từng tỉnh, thành phố. Cùng một loại cơ sở hoạt động trong cùng lĩnh vực nhưng lại mang các mã ngành hoạt động khác nhau.

Nguyên nhân của những tồn tại:

- Khâu tổ chức, phân công trách nhiệm cho các bộ phận ở từng địa phương chưa rõ ràng và chưa có cơ chế kiểm tra, nghiệm thu, giám sát cuối cùng trước khi tổ chức nhập tin.

- Nhiều bộ phận cùng tham gia với số cán bộ trình độ khác nhau nên nhận thức của mỗi người ở một số chỉ tiêu phân loại cơ sở cũng khác nhau.

- Chưa có hướng dẫn cụ thể cho mọi người để đưa ra qui định thống nhất chung cho khâu kiểm tra, ký mã.

- Tổ chức kiểm tra trong quá trình nhập tin chưa thường xuyên, chưa uốn nắn kịp thời các sai sót do các cán bộ nhập tin gây ra trong quá trình nhập tin phiếu điều tra.

5. Nghiệm thu các đơn vị nhập tin

TTTH Thống kê phối hợp với tổ thường trực của Ban chỉ đạo tổng điều tra của Tổng cục tổ chức nghiệm thu như sau:

- Nghiệm thu trực tiếp tại 7 tỉnh, thành phố (trong đó có Hà Nội, Tp. Hồ Chí Minh). Thời gian nghiệm thu trực tiếp tại các địa phương từ 1 đến 2 ngày, chủ yếu kiểm tra lỗi số học, lỗi logic, thực hiện truy vấn thông tin bất thường, tổng hợp nhanh một số chỉ tiêu để đối chiếu, kiểm tra chéo giữa các chỉ tiêu điều tra trong từng loại phiếu. Qua đó rút kinh nghiệm, đưa ra lộ trình nghiệm thu gián tiếp cho các tỉnh, thành phố còn lại.

- Các tỉnh, thành phố còn lại đã nghiệm thu qua các File dữ liệu nhập tin gửi về TTTHTK. Thời gian nghiệm thu là sau khi các địa phương đã gửi hết tệp dữ liệu nhập tin về, TTTHTK phối hợp với các vụ liên quan tổ chức phân công trách nhiệm nghiệm thu.

- Phương pháp nghiệm thu: tập trung cán bộ của các vụ liên quan trong thời gian khoảng 2 tháng tại phòng máy đào tạo của TTTHTK, phân công từng cán bộ kiểm tra, nghiệm thu từng loại phiếu cho từng địa phương với các qui định thống nhất về nội dung và phương pháp nghiệm thu.

- Kết quả nghiệm thu: cơ bản đã hoàn thiện một số chỉ tiêu phân loại cơ sở theo qui định chung thống nhất. Sửa các lỗi số học và lỗi logic ở các phiếu điều tra còn lại do các địa phương chưa hoàn thiện.

Tuy nhiên, một số tồn tại có thể rút ra trong khâu này là:

- Do khối lượng thông tin lớn, có quá nhiều các chỉ tiêu điều tra cần kiểm tra, do vậy các cán bộ nghiệp vụ chưa kiểm tra hết các quan hệ của các chỉ tiêu liên quan.

- Tổ chức nghiệp vụ tập trung nhưng chưa có cơ chế kiểm tra lại và kiểm soát cuối cùng đối với cán bộ tham gia nghiệp vụ, nên còn bỏ sót hoặc chưa thống nhất phương pháp kiểm tra.

- Chưa rà soát, đối chiếu số lượng các cơ sở đã điều tra với số cơ sở đã nhập tin vào máy cho từng địa phương do vậy chưa phát hiện được các địa phương gửi thiếu file dữ liệu đã nhập tin về Tổng cục.

- Xem xét kết quả tổng hợp nhanh một số chỉ tiêu của từng loại phiếu điều tra. Tuy nhiên do chưa đối chiếu, so sánh với kết quả nghiệp vụ từ tổng hợp thủ công của từng địa phương nên chưa phát hiện kịp thời sự chênh lệch giữa kết quả tổng hợp nhanh trên máy và tổng hợp thủ công.

6. Tổng hợp kết quả điều tra

- Phương pháp xử lý: xử lý, tổng hợp theo từng địa phương với nguồn số liệu của từng địa phương sau khi đã được làm sạch cùng với số liệu điều tra của Bộ Quốc phòng và Công an. Chuyển giao kết quả tổng hợp chung, tổng hợp từng loại phiếu điều tra của từng tỉnh, thành phố cho tổ thường trực ban chỉ đạo Tổng điều tra Trung ương để kiểm tra, xem xét, cho ý kiến để hoàn thiện số liệu lần cuối trước khi tổng hợp chính thức.

Qua kết quả tổng hợp này đã giúp tổ thường trực ban chỉ đạo Tổng điều tra Trung ương phát hiện những tồn tại trong quá trình thu thập và nhập tin, xử lý số liệu của từng địa phương. Đồng thời cũng phát hiện những điểm còn chưa thống nhất về mã hoá một số chỉ tiêu về phân loại cơ sở giữa các địa

phương, dẫn tới kết quả tổng hợp ở một số chỉ tiêu chưa hợp lý.

- Từ phát hiện những sai sót qua biểu tổng hợp, TTTHTK phối hợp với tổ thường trực, hoàn thiện số liệu lần cuối, tổng hợp kết quả chính thức và chuyển giao cho các địa phương trên đĩa CD ROM vào tháng 8/2003 tại hội nghị triển khai ở Huế. Qua đó các địa phương xem xét, đánh giá kết quả tổng hợp, phát hiện những bất hợp lý còn tồn tại trong số liệu điều tra của địa phương, trao đổi kịp thời với TTTHTK để hoàn thiện trước khi chuyển sang xây dựng các ấn phẩm điện tử (Ví dụ, qua đó phát hiện Long An gửi thiếu 1 cập nhập tin phiếu 3, Sơn La gửi thiếu 1 huyện nhập tin phiếu 4, Cao Bằng nhầm lẫn các tệp truyền tin).

7. Xây dựng cơ sở dữ liệu, đĩa CD ROM kết quả tổng điều tra

Cơ sở dữ liệu vi mô về số liệu điều tra gốc cho từng loại phiếu đã được xây dựng và được cài đặt trên mạng GSO_NET. Cơ sở dữ liệu vĩ mô về kết quả tổng hợp chung và kết quả tổng hợp cho từng loại phiếu cũng đã được xây dựng và cài đặt trên mạng GSO_NET. Đĩa CD ROM theo nội dung của 3 tập sách kết quả tổng điều tra bằng tiếng Việt và tiếng Anh đã được phát hành.

Bài học rút ra là cần phải có sự phối hợp chặt chẽ hơn giữa các đơn vị liên quan (như TTTHTK, tổ thường trực ban chỉ đạo Tổng điều tra Trung ương, nhà xuất bản Thống kê); các sản phẩm điện tử nên giới thiệu và hướng dẫn rộng rãi tới các đơn vị trong và ngoài ngành để mọi người có thể trực tiếp khai thác kết quả tổng điều tra từ các ấn phẩm điện tử. Cần tổ chức các lớp đào tạo sử dụng cho các cán bộ nghiệp vụ cách khai thác kết quả tổng điều tra từ CSDL ■