

关于货运统计的最短经由处理问题

岳 连 枚

(经济管理系)

摘 要

本文针对铁路货物运输统计中复杂、关键的经由处理问题,分析了通用程序未能投入使用的原因,对现行方案作了评价,提出了改进方案的设想。

计算机在管理中的应用,最初是从统计汇总开始,而且多数从货票统计着手。现在,全路各统计工厂都已实现了客货票精密统计电算化,正在继续向统计信息的网络化方向发展。因此,对现行的统计汇总实际进行总结、评价,以利于建立自动化统计信息系统,是十分必要的。

一、问题的提出

货物运输统计反映铁路运输的产品数量及质量。货物周转量是铁路运输产品之一,其计算方法为:

全路货物周转量 = $\sum$ (各批货物的重量 $\times$ 相应的运输距离)

运输距离是指货物发、到站间的计费里程,除特殊经由外,一般按最短经由计算。

我国铁路线路成网,有五千多个营业站,要保证在统计规定期限内,按线路和运输种类,分上、下行统计,信息量还很大,一个铁路局每月发生信息达数万笔以上,计算任意两个站间的最短经由十分复杂,其它统计报表比较简单,易于编制程序实现,唯独经由处理程序复杂而难以实现,这就成为货物运输统计的复杂、关键的问题。

有关最短经由的计算问题,已有各种通用的计算程序给以解决,不在此讨论。七十年代末,铁路统计初始应用计算机时,有人曾设计过全路精密统计通用程序系统,采用网络树和位势的计算方法,自动查找发、到站间经由轨迹和计算最短里程。当任意输入发、到站码,机器便可准确输出沿途经由轨迹和最短里程,做了一定贡献。该程序系统虽然调试通过,但因各种原因未投入使用。目前基本上仍采用仿照手工作业方案进行数据处理。本文就是在总结有关数据处理经验的基础上,分析了上述两种处理方案的利弊,探讨如何吸取各自优点进行最短经由的处理。

本文于1990年9月22日收到

二、通用程序未能运用的原因

统计处理分统计汇总和统计分析。通过汇总数据的汇总来掌握事物整体情况的过程，称为统计汇总；使用汇总的数字进行事物整体现象的分析，则称为统计分析。经由处理程序是统计汇总程序，它不同于数学公式简单的数据代入，而是大量的数据处理，受许多因素的制约，难以整齐划一，需要按照统计业务的要求，进行具体分析，设计各自的程序系统。

通用程序未采纳使用，究其原因，正如日本统计局铃木芳雄在“统计汇总的程序设计”一书中所述：“……在数据量很大，统计表复杂、繁多的情况下，通用的汇总系统在表现能力和处理时间上还存在一定问题，因而有时不得不放弃应用”。

通用程序对原有统计手工处理方案的优点分析不够，对待复杂的经路问题，采取全路统一处理，不仅使问题得不到简化处理，反而更为复杂化。表现在：

(1) 查询范围太广

由于全路统一计算周转量，是把每笔信息从发站一直计算至到站，这样需要把全路站码、计费里程全部输入机器，占用机器很大空间，查询范围太广。

(2) 经路越长，处理愈复杂

由于全路计算货物运送吨数，需要详细标出每笔信息的转线轨迹，找出经由每条线路的中转站，每通过环状线网，便要对最短经路作判断，路程愈长，通过的环状网越多，则经由愈复杂，查询时间延长。

(3) 特殊情况难统一

由于各个铁路局线网环境不同，货物品类不同，如东北发送石油，不按最短经由统计，则使通用程序不论南、北方局，一律作是否属特殊经由判断。这样将全路各种特殊情况汇集统一处理，使问题复杂化。

由于上述原因，使程序运行时间过长，加上当时国产 C4 机运行不稳定，而未被运用。

三、现行方案及其评价

现行方案是仿照原统计手工处理系统。其处理方法为：

1. 以字典处理代替计算处理

现行方案是按铁路局汇总运送吨数，不是把每笔信息从发站一直计算至到站，而只计算到分界口为止，出局界由他局处理。其汇总顺序如下：

区间周转量 = $\sum$ 各批货物吨数 $\times$ 区间里程

区段周转量 = $\sum$ 所在区间周转量

铁路局周转量 = $\sum$ 局管辖区段周转量

全路周转量 = $\sum$ 铁路局周转量

统计周转量，先汇总经由区间的四种运输种别的货物吨数，然后统一乘以区间里程，求得经由该区间的货物周转量，最后逐级汇总为全路周转量。手工方案把报告期的区间里程看作是固定信息，是一个常量，而货物运送吨数则看作为变动信息。所以，周转量的统计处理实质上是运送吨数的汇总处理。

运送吨数的汇总是分上、下行方向之别的，一笔信息是汇总在线段的上行还是下行，它是根据发、到站间的最短经由判定后而确定的。而最短经由是依据计算里程而定。因此，现行方案把经由处理分两步进行：第一步，将计算里程确定的经由编入字典。第二步，查询字典判定上、下行别。程序运行时，直接借助字典作为处理工具。

2. 按照分歧站编排经由字典

所谓分歧站是指对应于环状线上某站，并指明该环状线不同方向的最短经路的分界点。见图一，以绥（化）佳（木斯）牡（丹江）哈（尔滨）环线为例，

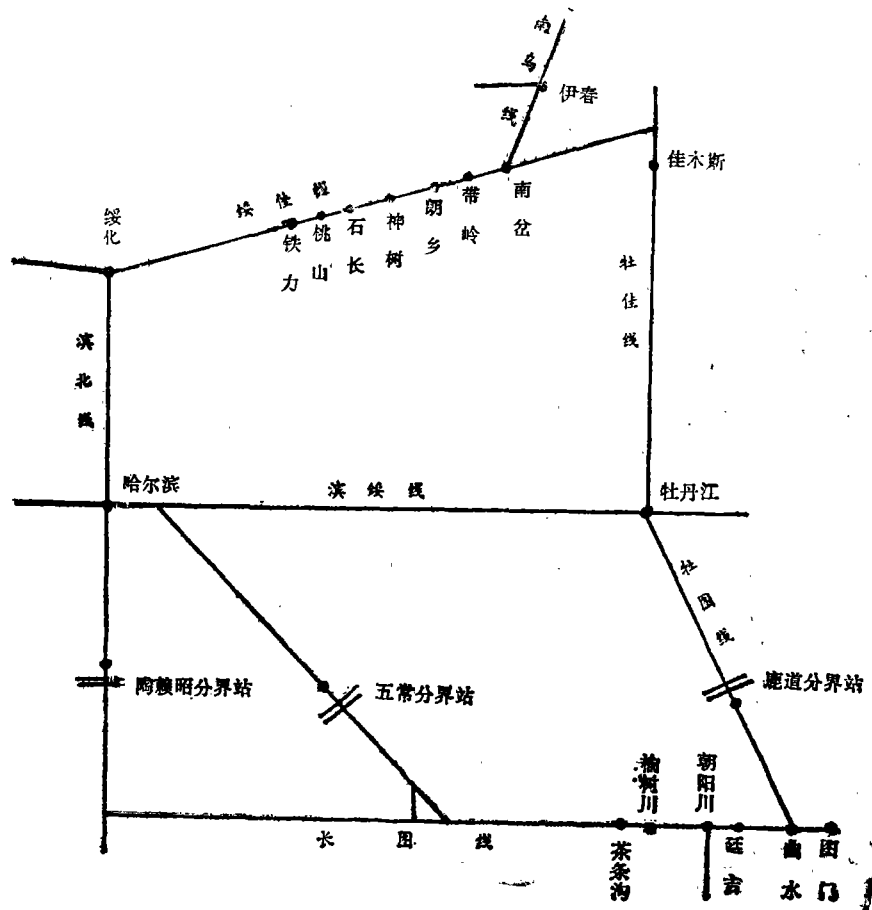


图 一

桃山站对应于牡丹江站，牡丹江到桃山以东各站经由牡佳线为近，到桃山（包括本站）以西各站经由滨绥线为近，桃山站就是两个方向的分界点，即分歧站。

发、到线经路可以归纳为如下三种情况：

(1) 无分歧站

绥佳线发货往滨北线，全部经由绥化站，全线设有分界点。

(2) 单一分歧站

绥佳线发货往牡图线，全线以桃山站为分界点，只有一个分界点。

(3) 多个分歧站

绥佳线发货往长图线，有榆树川、朝阳川、延吉、曲水……等多个分界点。神树站发往长图线货物，到站以曲水站为分界点，即到曲水以东（包括曲水）各站，经由鹿道口为近，到曲水以西各站，经由五常口为近；朗乡站发往长图线货物，到站以延吉站为分界点，即到延吉以东（包括延吉）各站，经由鹿道口为近，到……，从而，可编排如下经由字典：

营 业 站		经由鹿道口的分歧站
神	树	曲 水
朗	乡	延 吉
带	岭	朝 阳 川
南	岔	榆 树 川

3. 查字典判定上、下行

程序运行时，分别三种情况进行处理。若汇总由带岭发往曲水的货物吨数，查经由字典得知：带岭发往长图线的货物，到站以朝阳川为分歧点，该例到站曲水在朝阳川以东，则可确定这笔信息经由鹿道口出局，其转线则为牡丹线，转线站则是鹿道口和牡丹江。连接线为牡丹线，转线站则是牡丹江和佳木斯，最后是发线——绥佳线，所以佳木斯站为发线上的到站。在发线——绥佳线上，发站是带岭，到站为佳木斯，该批货物当然汇总在下行方向。

4. 寻找归结点

编排字典时，尽量合并相同经由，例如把处于非环状线上的货流归纳到环状线的出、入口。见图一，南乌线各站发、到货流均可归纳为南岔站发、到。这样，不必再编南乌线到各线的经由字典，直接借用南岔站到各线的经由。又如绥佳线各站进关货物因其经由与哈尔滨站相同，则比照哈尔滨站进关的经由。对进关货物除少数到京承线等外，一概都经山海关口入关，所以可归纳为山海关发往关内各站的经由。

5. 直通货物只标明出、入局分界口代号

由于各局汇总本局管辖范围内的发送吨数和接运外局的运送量，发送局只向他局提供经由分界口的资料，各分界口间最短经路是固定的，所以，只要标明出、入局界代号，即可确定直通货物的全程经路，不必详细标出沿途中转转线轨迹。如南岔——兰州局的武威南站，由于排除到北京局的经由，可归纳为南岔——山海关——武威南，南岔进关经陶赖昭口，便可确定为陶赖昭出口，再确定到局分界口，见图二。

由山海关发往兰州局的货物，除到达陇海线的唐家堡——天水间各站，经由天水分界口外，其余都归纳到石咀山分界口，从而确定经由为：山海关——石咀山——武威南。因为出分界口陶赖昭——入分界口石咀山的最短经由是固定的，这样不必计算里程，便可确定经由。

现行方案这些处理方法，体现了如下优点：

(1) 节省计算里程时间。因直接查询计算里程的结果——经由字典，所以，节省计算、查找最短里程的程序运行时间。

(2) 化繁为简。由于处理时可利用本局线网条件，寻找归结点，把大量随机产生的发、到站归结到固定的归结点，又因归结点到环状线经路分歧站固定单一，从而便于找到发、到局分界口，再利用各局分界口经由固定的条件，确定全路经路。使远程货物的复杂经

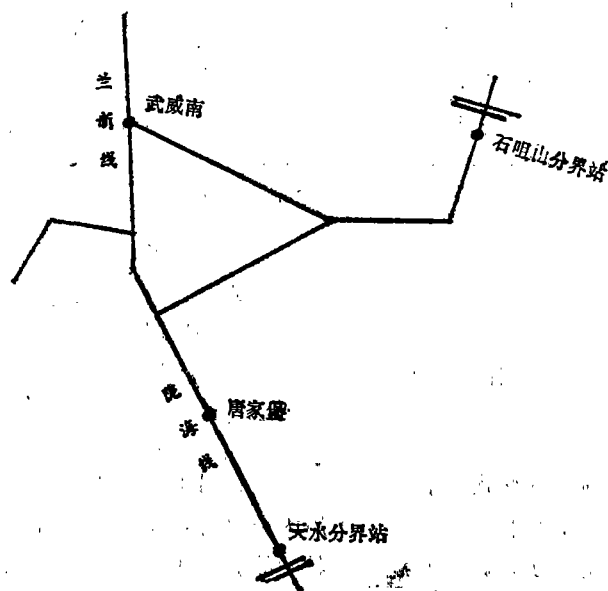


图 二

由处理得以简化。

(3) 节省存贮空间。由于大量信息经过归纳合并，变随机发生经由为固定经由，则可用少量字典代替繁多的计费里程，从而节省机器存贮空间。

然而，现行方案在处理中，也有其不足之处，主要是：

计算工作量大。编排经由字典要靠手工计算里程，遇到路网密集地区，大环套小环，发、到线间可能有多个分歧站，需要计算里程，反复对比才能编成字典，工作量很大，而且手工计算里程，编排经由字典，供机器查找，犹如“背水喂洗衣机”，不能充分发挥机器作用，还难以保证经由字典质量，一旦错编，由于数据处理重复作业，月计年累误差大，会直接影响货物周转量的统计质量。

程序系统难以维护。上述归纳合并相同经由，是一种凭经验的工作，遇到新的营业站投入使用，形成新的环状线网，或者铁路局界变迁，使程序难以在短时间内得到维护，有时则要推倒重编。

四、改进方案的探讨

为了充分发挥通用程序和现行方案各自的优点，更好地进行货物运输经由处理。笔者认为可按具体情况，分别进行处理。

(一) 直通远方局的货物经由处理。

直通远方局货物经由仍沿用现行方案，即手工编排字典处理。因为现行方案可寻找归结点，由归结点为起点计算最短经由，使经由字典编排简单易行。各铁路局也都有可能根据本局情况，找到归结点，同时多年积累了现成的经路资料，有利于手工编排字典，此外采用字典处理，可以节省机器运行时间和存贮空间，从而适宜车站微机运行，可为进一步实现“货票不进厂”等创造条件。

（二）直通邻近局的货物经由处理。

这种货物经由应采用靠通用程序提供资料、人工编排字典的处理方法。因为直通邻近局货物只需标明出、入局分界口代码，适宜用字典确定经由。但是由于发、到站相距较近，发、到线间经由分歧站多而不固定，计算里程工作量大，则以通用程序输出各条线路间发、到货物的经由轨迹和最短里程，为手工编排多经由字典提供资料。这样，把编排字典与查询字典分开进行，可节省程序运行时间和空间，避免通用程序的缺陷。

（三）管内货物经由处理

直接采用通用程序处理方案。即机器计算每笔信息的计费里程，自动查找经路，标出详细的沿途经由转线轨迹。因为发、到站均在管内，大量信息随机发生，而且按铁路局汇总运送吨数，也需要沿线转线轨迹，因此，不宜编排经由字典。又因发、到站均在局管辖范围内，只需输入管内站名和计费里程，占用机器空间不大，查询范围也大为缩小，查出来的转线轨迹，还可用来解决货流图枢纽分解的难题。

总之，货物运输统计经由的处理，需根据各铁路局的具体情况，采取各自简便的处理方案，才能节省机器存贮空间和运行时间，进一步与财务收入部门、各货运站共用通用程序，达到数据资源共享，加速统计信息自动化的进程。

参 考 文 献

〔日〕铃木芳雄著·崔晓光译·统计汇总的程序设计·北京：中国统计出版社，1985，2