ABSTRACT

Intercity traffic in Bell Canada is carried on an intertoll switched network with analog and digital toll switching machines subtending local switching networks in each major population center. The network is provisioned from a forecast of intertoll trunks interconnecting switching machines between major community of interest locations according to a fixed hierarchial routing plan.

A growth rate must be established for offered traffic to each interconnecting trunk group for input to the overall intertoll forecast process. Formerly this growth rate was derived from historical CCS usage of first route offered traffic. Year over year comparisons became difficult with rehoming and detolling due to changes in network architecture.

This paper describes a more accurate process termed the Growth Rate Forecasting System (GRFS) that has been introduced in the company to derive trunk group growth rates from historical point to point calling data independent of network architecture change.

INTRODUCTION

The Bell Canada intertoll network services 9 million phones in an area 3400 km in width by 2000 km in depth situated in the central portion of the country interfacing with networks in the remainder of Canada, the United States of America and other countries throughout the world. The network comprises 100 toll switching centers of varying types from step-by-step and cross-bar through to stored program control analog and digital switching machines. These are interconnected with final trunk groups in a fixed hierarchial arrangement from lower class 4 switching centers to primary class 3, sectional class 2 and finally the regional class 1 switching center in Montreal. The Montreal regional center is connected with other regional centers in Canada and the USA with overseas calling provided via Teleglobe Canada. Approximately 1000 high usage trunks carry intercity Traffic between toll switching machines where a high community of interest exists. Overflow in peak periods from these groups is carried on final groups in ascending order from class 4 to class 1 machines. Individual phones gain access to this intertoll network via class 5 switching machines and toll connecting trunks in local networks connected to toll switching machines throughout the territory served.

Provisioning on the intertoll network originates with the forecast of intertoll trunks on the final and high usage (HU) trunk groups. These trunk quantities manifest themselves as terminations on toll switching machines along with local network toll connecting trunks. Local network forecasting is handled on the basis of CCS/main station. (1,2) Intertoll network forecasting is done on the basis of forecasting traffic usage first offered to each individual trunk group in the current year and then growing each amount by its own individual growth rate. This Traffic is then combined on trunk groups according to the hierarchial routing rules and network architectural structure planned for each period of the forecast. The CCS of traffic usage so allocated to each trunk group in the forecast manifests itself in terms of traffic trunks required year by year for that group according to criteria set for maximum call blocking and well known principles of traffic engineering.

This paper deals with the most critical part of the above intertoll forecasting process which is to establish a growth rate for each parcel of traffic offered to individual trunk groups on the intertoll network.
THE PROBLEM

Each end office in the local network attached to a toll switching machine originates a pattern of intertoll calling to the world dependent on its community of interest factors. (3) An allocation of toll calling to high usage trunk groups emanating from the toll switching machine thus is dependent on both the number of end offices homed on the toll machine and the pattern of their calling. If the network architecture did not change from year to year, that is, if the number of end offices and their homing remained constant on a fixed number of toll machines, forecasting a growth rate for intertoll trunks would be a very simple process. The offered traffic to each could be established directly from historical trends using various algorithms for future projection.

In the static toll network configuration above, traditional CCS usage, peg count and overflow readings are taken during the peak period to determine network cluster busy hour offered traffic to each HU trunk group (and first route offered to the final) in the current basing period. Assuming the network architecture homing arrangements remain unchanged, the growth rate forecast for each trunk group for input to the intertoll trunk forecasting process can be obtained simply by applying various forecasting techniques to historical trend data of first route offered CCS to that particular trunk group, eg. HAMILTON 1 to TORONTO 3.

Unfortunately, toll switching machines do not have unlimited capacity. When additional end office loading increases to the point where second and third toll machines are required in one location, end offices are rehomed to growth machines. Class 4 toll machines may be rehomed to new class 3 growth machines. The result is a complete change in the first route offered traffic to interconnecting high usage and final groups. This change in architecture, if extensive, destroys the relationship of current trunk usage values to historical trends preventing an accurate forecast of growth rates for this trunk group based on usage readings only. If minor changes in homing occur on a lower class machine, it may be possible to accommodate it by manual adjustment to usage readings. When several changes occur simultaneously it is virtually impossible to establish a relationship which will produce an accurate comparison.

In Bell Canada, network architecture changes were minor until recent years when rapid toll growth caused initial cross-bar machines to exhaust and be replaced by stored program machines. This was followed by the installation of new technology digital machines. Although the intertoll forecasting process involved a point to point disposition of call analysis to be correlated with usage measurements during a specific period each year, it was still difficult to trend and forecast usage results from year to year.

THE SOLUTION

To overcome discontinuities in usage trends brought about by rehoming, detolling, addition of new toll machines, etc., it was realized that historical data more basic to the origin and destination of traffic and independent of the network structure was required. If this data could be gathered between community of interest calling centers and forecasted then a transformation might be made to determine how this would result in growth of first route offered traffic from usage measurements on the current network structure.

All intertoll traffic calling is recovered on Automatic Message Accounting (AMA) systems associated with each toll switching machine. An originating subscriber's end office number, the start and finish times of the call, the called number, type of call, time and date are recorded. This appeared to be an excellent source of toll data, representing a matrix of ABD CCS usage between community centers and independent of the network. Although the data was voluminous and varied greatly from low to high community of interest centers a variable sampling technique was required to originate a statistically sound representative value of CCS usage in this fashion. A program was written to derive this information daily from 8 AM to 11 PM, compute monthly average business day (ABD) values, and store them in a data base representing a matrix of ABD CCS usage between called and calling toll switching centers for each month of the year. This was computed and stored from six consecutive historical years of AMA data previous to the first year of the forecast.

Fig. 2 Intertoll calling allocation to high usage and final trunk groups from specified end offices in a fixed configuration.

Fig. 3 Network architecture change from that shown in fig. 2 disrupting calling pattern trends on Hamilton 1 to Toronto 3 trunk group.
work must have routing instructions in its data base for forwarding all calls received such that each call will be allocated to its proper high center throughout the six year period (or portion thereof for new offices installed in that time frame). Unfortunately this is not the case as rehoming to new toll machines must take place to accommodate growth. This problem was solved by assuming end offices were homed throughout the six years on the toll center they were homed in the base year. This then gave a calling growth trend from the base year into the past which is exactly the historical data required to forecast from the network configuration of the base year into the future!

Each toll switching machine on the intertoll network must have routing instructions in its data base for forwarding all calls received such that each call will be allocated to its proper high usage (or final) trunk group as first choice to reach the called switching machine. A centralized data base was set up to administer routing updates input to each machine. For the base year preceding the forecast, it was possible to access this data base and obtain both end-office/toll center association data and the first choice route for traffic offered from end office to end office over the network between toll centers. Again, the same principle was applied and the trunk group configuration existing in the base year for first route and overflow traffic was assumed to be that which existed for the previous six years.

Having historical trends of CCS calling usage from point to point data and their allocation to trunk groups on a first route offered basis in the base year it was now possible to associate these with actual offered traffic in the base year developed from measured values of usage, peak count and overflow. Furthermore, a growth rate forecast of CCS calling data could be deemed the growth rate forecast of first route offered traffic derived from these measured usage values. The only problem remaining was to develop programs, utilizing established techniques, to forecast the CCS calling data and the overall objectives would be attained.

The foregoing may be expressed as follows for each pair of calling and called switching centers:

\[ q_t = \frac{d_n}{N} \]

- \( q_t \) - average day calling CCS for month t (0 < t < 72)
- \( d_n \) - number of days/month for data collection from AMA system
- \( N \) - monthly ABO calling CCS offered to London - North Bay trunk group
- \( f_t \) - average first route offered traffic derived from measured usage on the interconnecting trunk group(s) during busy hour
- \( k \) - proportionality constant assumed for point to point message vs first route offered traffic

assum ing

\[ h_n = \frac{d_n}{d_{n-1}} \]

- \( h_n \) - CCS calling in busy hour during day n

that is

\[ \frac{h_n}{h_{n-1}} = \frac{d_n}{d_{n-1}} \]

we have

\[ dq = k \frac{df}{dt} \]

Thus a growth rate forecast of \( q_t \) will yield the rate of growth forecast for \( f_t \) which is input to the intertoll trunk forecast system.

The above calling data \( q_t \) stored in the Growth Rate Forecast System (GRFS) data base may be expressed as matrices \( M_t \) for all monthly CCS historical calling data between toll switching centers x and y in a network configuration defined for the base year.

**FORECASTING A GROWTH RATE**

Having established a history file of calling data
for each trunk group which could be plotted from the database as shown in Fig. 4, it now remained to devise a mechanized forecasting module which would screen this data, make any transformations required, handle extreme values, prepare required transformations, apply various forecasting modules, select the appropriate model and compute future forecasted values. An algorithm of GRFS termed "FORMANAL" was used for this purpose.

**THE FORMANAL ALGORITHM**

1. **STAGE I**
   - **PRELIMINARY DATA ANALYSIS**
   - **SEASONAL TREATMENT**
   - **TRANSFORMATION SELECTION**

2. **STAGE II**
   - **MODEL-BUILDING AND SELECTION**

3. **STAGE III**
   - **FORECASTING**

Raw seasonal indices are computed by comparing actual history to the moving average. Two types are employed as follows:

\[ R_t^{(\text{ratio})} = \frac{q_t}{M_t} \]

\[ A_t^{(\text{additive})} = q_t - \hat{R}_t \]

The test for seasonality is based on a one way analysis-of-variance of the raw indices comparing the between months variation to the within months variation. The resulting ratio is termed the "F value" and is employed in the "Stable Seasonality Test". The computed F value is compared against a critical value of 1.5. Values greater than 1.5 indicate the presence of monthly seasonality. The F test is performed on both ratio and additive indices. If both indicate the presence of seasonality then adjusted seasonality factors are computed to deseasonalize the history data in order to derive the data trend and determine whether or not a transformation is required.

\[ \tilde{R}^{(\text{ratio})}_k = c_1 \bar{n} + c_2 \sum_{j=1}^{n_k} R_{kj}/n_k \]

\[ \tilde{A}^{(\text{additive})}_k = \bar{n} - c_2 \sum_{j=1}^{n_k} A_{kj}/n_k \]

The ratio or additive approach is then selected based on the best fit to the actuals (minimum sum of squared errors). Once chosen, the adjusted indices are used to determine deseasonalized \( q_t(D) \) calling data values as follows;

\[ q_t(D) = q_t^{(\text{ratio})} \text{ or } q_t^{(\text{additive})} = q_t - \hat{A}_t \]

The \( q(D) \) deseasonalized calling value history data are checked next with least square fits of linear and exponential trend curves to determine if a log transformation is required to convert exponential growth patterns to a linear format before inputting calling data to various models for forecasting.

Assuming \( q_t(D) = a + bt + e_t \)

the residual sum of squares, \( RSS_0 = e_t^2 \) is found. Similarly for \( \ln q_t(D) \)

\[ RSS_1 = e_t^2 \]

A check is made next to test for data seasonality by employing a 12-term centred moving average which produces values appropriate for all months of history except the first 6 and last 6:

\[ M_t = \frac{1}{24} (q_{t-6} + 2(q_{t-5} + q_{t-4} + q_{t-3}) + q_{t-6}) \]

\( M_t \) - moving average value of calling data corresponding to \( q_t \) actual
RSSL = I(exp(RSS)) is determined

Transforming back to original terms yields a residual sum of squares from the estimated exponential curve:

RSSL = I(exp(RSS))

If RSSL > RSSL, a log transformation is selected, otherwise original q_t calling data is used.

The treatment of extreme (outlier) q_t values is the last step taken in the preparation of history data for modelling. This is carried out on deseasonalized and, if required, transformed q_t values. A 13 term Henderson curve(4) is used for this purpose. It is similar to a moving average approach except specific weighting is required. Residual E_t's are calculated as follows:

E_t = q_t(D) - H_t

H_t - Henderson curve value for month t

The standard deviation, s, is estimated by

s = \sqrt{\frac{\sum (E_t - \bar{E})^2}{N}}

If |E_t| < 1.5s, no adjustment is made
If 1.5s < |E_t| < 2.5s an adjustment is made to reduce E_t in accordance with the following:

E_t' = E_t (2.5 - |E_t|/s)

If |E_t| > 2.5s then E_t is made equal to zero

The above approach is an adaptation of the "Graduated Treatment of Extremes"(4).

Having determined if seasonality is present, whether a log transformation is required and having tagged and adjusted outlier values in history calling data, the application of various forecasting models to this data can take place.

Three basic models are used; regression with dummy monthly seasonal variables(5) auto regression(6) and exponential smoothing. Transformation, where required, and extreme value treatment on outlying values of q_t apply on history data for all three. Only the last, exponential smoothing, works against deseasonalized history data.

The three models may be written as follows:

(a) regression

q_t = \beta_0 + \beta_1T + \beta_1D_1 + \ldots + \beta_11D_{11} + \epsilon_t

(b) auto regression

q_t = \phi_0 + \phi_1q_{t-1} + \phi_2q_{t-2} + \ldots + \phi_pq_{t-p} + \epsilon_t

(c) exponential smoothing

q_t = \alpha q_{t-1} + \alpha(1-\alpha)q_{t-2} + \alpha(1-\alpha)^2q_{t-3} + \ldots

T - independent variable, months

D_t - monthly dummy variables
q_t - actual value of calling
\epsilon_t - error component in month t

Each model is run to first determine the best fit values to history data and second to forecast future values of q_t.

Model selection is based first on the minimal standard error between actual (q_t) and fitted (\hat{q}_t) historical values of calling data:

SE = \sqrt{\frac{\sum (q_t - \hat{q}_t)^2}{v - p}}

v - number of fitted values, q_t

The above approach is an adaptation of the "Graduated Treatment of Extremes"(4).

Along with the production of fitted values to historical data and a future forecast of fitted values of CCS calling, the GRFS system also performs a monitoring function by running a forecast of q_t values for the current base year (from 5 previous years of historical data) which may be compared with actual q_t values in the base year to determine the validity of the forecast in this period as a measure of its trustworthiness in the future.

A typical plot illustrating the GRFS system output is shown below with fitted (X), base year monitoring forecast (M) and future forecast (F) values of q_t.
Fig. 6 Historical, fitted and forecast calling CCS offered to the London-North Bay high usage trunk group.

Growth rate forecasts of CCS calling first route offered to this trunk group from the base year 1978 forward are as follows:

<table>
<thead>
<tr>
<th>Year</th>
<th>Forecast Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>8.44%</td>
</tr>
<tr>
<td>1979</td>
<td>8.42%</td>
</tr>
<tr>
<td>1980</td>
<td>7.85%</td>
</tr>
</tbody>
</table>

Diagnostics associated with this forecast are summarized below:

<table>
<thead>
<tr>
<th>Method</th>
<th>Avg. Error</th>
<th>Abs. % Error</th>
<th>Fit Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>156.70</td>
<td>8.70</td>
<td>86.62</td>
</tr>
<tr>
<td>Auto regression</td>
<td>173.48</td>
<td>9.23</td>
<td>74.86</td>
</tr>
<tr>
<td>Exp. Smoothing</td>
<td>181.17</td>
<td>9.06</td>
<td>80.32</td>
</tr>
<tr>
<td>Raw Data Screen Code</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transformation Index</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seasonality Statistic</td>
<td>1.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contamination Index</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Examination of the above indicates that regression (with dummy seasonal variables) provides the best overall fit according to the Standard Error, Average Absolute % Error and the Fit Index. The Raw Data Screen Code (range 1-5) indicates no missing history data and the Contamination Index being less than 0.2 indicates that the data had relatively few extreme values. The Seasonality Statistic (F value) indicates seasonality is present in the data as it exceeds the critical 1.5 value. Also, a log transformation was employed as the RSSo/RSSl Transformation Index is > 1.

The forecast of calling CCS in the base year (months 60 to 72 from 5 year history data) shown by M values on the plot may be compared with actual values of \( q_t \) in this period. A comparison of the forecast vs actual CCS growth in this monitoring time frame gives the following values:

<table>
<thead>
<tr>
<th>Year</th>
<th>Actual</th>
<th>Forecast</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977-78</td>
<td>12.96</td>
<td>13.94</td>
</tr>
</tbody>
</table>

GRFS SYSTEM TRIAL

The GRFS system was run experimentally in 1977 as described in the foregoing. During 1978 the regression with dummy seasonal variables was revised to include monthly wage levels in the region containing the called and calling toll centers in order to bring in the effect of an economic variable on the forecast. Sample results of the forecast from this revised model for the monitoring period 1977-78 are summarized below.

<table>
<thead>
<tr>
<th>Trunk Group</th>
<th>Actual</th>
<th>Forecast</th>
</tr>
</thead>
<tbody>
<tr>
<td>London 1 - Oshawa</td>
<td>16.5</td>
<td>15.4</td>
</tr>
<tr>
<td>Brockville - Toronto 3</td>
<td>12.8</td>
<td>14.3</td>
</tr>
<tr>
<td>Belleville - Peterborough</td>
<td>10.9</td>
<td>11.1</td>
</tr>
<tr>
<td>Brampton - Kitchener</td>
<td>11.1</td>
<td>11.3</td>
</tr>
<tr>
<td>Brantford - Hamilton 1</td>
<td>10.3</td>
<td>10.8</td>
</tr>
<tr>
<td>Brantford - London 1</td>
<td>10.6</td>
<td>10.6</td>
</tr>
<tr>
<td>Beaverton - Oshawa</td>
<td>12.0</td>
<td>13.3</td>
</tr>
<tr>
<td>Cornwall - Ottawa 1</td>
<td>8.3</td>
<td>8.6</td>
</tr>
<tr>
<td>Port Erie - London 2</td>
<td>11.1</td>
<td>11.7</td>
</tr>
<tr>
<td>Port Erie - Niagara Falls</td>
<td>9.9</td>
<td>8.2</td>
</tr>
<tr>
<td>Guelph - London 1</td>
<td>12.2</td>
<td>13.5</td>
</tr>
<tr>
<td>Kitchener - London 2</td>
<td>10.6</td>
<td>9.8</td>
</tr>
</tbody>
</table>

Results of the trial to date look promising. Forecast figures produced in the monitoring period are very close to those actually incurred. More time is required, however, to evaluate current forecasts against future actual values.
Another application of GRFS to be considered is the use of the head-end point-to-point CCS calling data forecast algorithm as an input to a dynamic routing(7) forecasting system in place of inputting to a fixed hierarchial scheme as described in this paper. In dynamic routing, traffic is routed from the originating office to the terminating office over one of many tandem routes that are available depending on the routing algorithm applied, e.g. - least busy in the time frame of the call, least cost, etc. As the routing varies with each call, a system such as GRFS, which derives its growth rates independent of call routing, is a natural for this purpose providing each switching machine has a method such as AMA for recording each call.

CONCLUSION

The GRFS system is a practical tool to establish growth rates for first route offered traffic to intertoll trunk groups. This information is required for input to most intertoll trunk forecasting systems such as Bell Canada's Trunk Planning and Forecasting System program which formerly derived these growth rates from trunk group usage measurements.

Practical experience has shown improved forecasting on the intertoll network using the Growth Rate Forecasting System techniques. This has resulted in manpower and dollar savings with better service to the customer as the intertoll forecast determines the quantity of facilities, switching capacity and extent of future change to be provided in the downstream provisioning activities. Any error in growth rate has a direct bearing on this process.

ACKNOWLEDGEMENT

The author would like to express his thanks to A. Buxton, HQ Corporate Forecasts and G. Berdebes, Corporate Systems Organization, Bell Canada, for their assistance on this project.

REFERENCES


(2) N.D. Blair, AT & T Basking Ridge N.J. "Forecasting Telephone Traffic in the Bell System" I.T.C. 8th Congress, Melbourne - 1976

(3) A.W. Dunstan, Telecom Australia, Brisbane "A Mathematical Model of Telephone Traffic Dispersion in some Australian Metropolitan Networks" Australian Telecommunication Research - Vol II #1, 1977


(7) Szbicki, E. and Lavigne, M. Bell Canada, Ottawa "The Introduction of an Advanced Routing System into Local Digital Networks and its Impact on the Networks' Economy, Reliability and Grade of Service" I.T.S. Paris - 1979