Transactions on the Built Environment vol 34, © 1998 WIT Press, www.witpress.com, ISSN 1743-3509

An analytical model for the simultaneous calculation of capacity of lines, junctions and station tracks

Livio Florio & Lorenzo Mussone
DSTM, Politecnico di Milano, Italy
Email: mussone@mail.polimi.it

Abstract

The calculation of capacity of railway lines is performed today by the railway management by using empirical or analytical methods. The paper aims at demonstrating the validity of the above-mentioned methods for each railway scheme upon their integration with analytical constraints for junction and station trucks capacity, components of a given ailway scheme. Capacity is calculated as solution of a problem of a constrained optimum.
An application of the proposed method is carried out for a real case of a medium complexity railway scheme. This allows us to verify immediately its validity and to evaluate planning alternatives of management.

1 Introduction

The increased interest for railway systems has lead, in recent times, to the continuous search of analytical methods and simulations to determine the capacity of railway lines. At the same time, methods have been studied for the capacity of the railway junctions and for station tracks, although all the components are to be considered as parts of a more general railway system or of a system assimilable to it [9](Florio and Mussone 1995). The capacity in any point of a "railway system" can be expressed as the maximum number of trains that can occupy it in a set time interval,

T. It must be clarified that, in practice, every "railway system" is identified for each scheme component (line, station and junction track) with points where trains start or end their movements, called singular points. From the analytical point of view, therefore, in a generic point as above described, it must be determined the value of

$$L = \frac{T}{\bar{h}} \qquad 1)$$

where L = capacity (number of trains), T = reference period, $\bar{h}$ = average interval between trains.

For T (the reference interval), it is common to consider an hour (hourly capacity) although other values can be used: like the day (daily capacity) or the peak hours in the management (capacity in the peak hours). The average interval between trains ($\bar{h}$) corresponds to the headway in the classical theory of road flow, but in the railway case the average interval assumes different meanings according to the considered scheme component. More exactly:

a) in the case of the lines, it is indeed possible to speak of average interval between trains, equal to the one a hypothetical observer would measure standing at the sides of the track and seeing them pass;

b) for station tracks, the average interval between two following occupations of the same track must be measured, no matter if the trains must stop or not;

c) in the case of the nodes, the average interval between two following trains, departing from the same starting or ending point of the paths, must be calculated; these trains engage a small system made by only one server (the interested path) with more users (the trains) queuing up on paths incompatible with the first one. In these terms, the node constitutes a bottleneck in the railway flow.

Considering what has been mentioned above, it can be realised that the value of L is different in any point of the railway system as well as in any other system to it assimilable. For capacity of a railway system in T, it must be intended the whole number of trains that occupy its components (line and station tracks and complex nodes) and, therefore, the result of 1) in each point must be mutually harmonised in a way that, for instance, trains on the line can enter the station and viceversa.

The calculation of 1) is further complicated by frequent stochastic arrivals or casual passages of irregular (not on time) trains that have different effects on each part of the railway system and that causes an important variability of $\bar{h}$.

This paper is aimed to define the search on the subject and to propose new evaluations based on the "traffic vectors" (number of trains, for each category in a predetermined time interval, that affects one of the

components above mentioned). It is divided into four parts:
a) critical revision of the methods for the calculation of the line capacity (UIC and DB);
b) determination of the calculation methods of the capacity of junctions and station tracks;
c) definition of the traffic vectors for line, junction and station track according to the capacity of each component and according to predetermined hypotheses on the subdivision of traffic;
d) definition of the rules that a timetable must respect concerning both the capacity of the scheme components and the service quality, that is the maximum regularity of circulation and consequent minimisation of delays.

In the end, an application to a real case of the exposed concepts allows us to verify immediately the proposed procedure. It must be preliminarily clarified that the applicability of the exposed concepts is not related to the number of system components, so that no methodological distinction must be made with respect to the type of line (simple or double track), the composition of the station plants or to the complexity of the junctions present in more parts of the considered system.

2 Methods to calculate line capacity

Omitting the formulations of calculation of static capacity, which are actually old and are already being abandoned by the railway Administrations, the methods commonly used to calculate capacity of lines can be divided in two types: a)methods based on probabilistic formulas aimed at individualising the average interval of circulation on each line; b) methods using simulation procedures .

a)Among the first ones, the best known is called UIC (Union International des Chemins de Fer), the name of corporate body that recommends its adoption in his fiche [9](UIC, 1978). It is based on the calculation, for each line track, of the sum of three values t_{fm}, t_r and t_{zu}, respectively the average interval of minimum distance between trains (in minutes), the amplification margin forseen in the timetable (minutes), the extra time to be added in order to reckon with the number of stations on the part of line under exam (min); they must be determined in any single case according to the specific composition of the traffic. It is evident that the sum of the three variables, valued stochastically, is the $\bar{h}$ in 1). As an alternative, for the calculation of the capacity of daily traffic the method used by Deutsche Bundesbahn [1][2](DB, 1974; 1979) is also based on a formula similar to 1), in which however for T is assumed the value 435 and the value of $\bar{h}$ is obtained by a weighted average of all possible successions between trains on line. In equal conditions, the capacity values of the lines, obtained by the DB method, result sensitively inferior to those obtainable with the UIC method. This is mainly due to the high probability of obtaining a value over the average value assumed for traffic in the peak hours (1%) that, actually, limits capacity by a great deal. It must be however considered that the two analytical formulations more widely used are based on formula 1) differentiating from one another only for the value to give to T (any value in the UIC formulation , 435 min. in DB formulation) and how the $\bar{h}$ value is calculated.

b) It is only left to mention shortly the methods of simulation that allow the calculation of the capacity obtaining in an indirect way values of T and of $\bar{h}$, basing on a big number of hypothesised situations [7](Petersen and Taylor, 1982). Their result, approaching to real situations, could be similar to the one obtainable with the UIC or DB method according to the hypotheses assumed for the simulation. It must be said that the simulation methods are very flexible. They can adapt themselves to consider all the real situations of the distance on line, track schemes and stochastics in the management. If, on the one hand, they don't set limits in the application, on the other hand they allow to affirm that their validity is always defined within the specificity of the examined situations, as it is possible to obtain a different capacity value for each particular hypothesised situation.

At the end of this brief excursus, it can be said that the search of the average interval between trains, travelling on the same track of line, is the only element in common among the three methods above illustrated.

It must also be considered that today reduced intervals are possible for the distancing of trains; they allow the assimilation of interval between trains to that of public transports based on car-following. Such methods can be based on the very precise control and location of each single unit (Automatic Train Control), also using satellitar systems (GPS or Ground Control Position) if necessary. Even using such powerful instruments, the stochastics of irregular trains in respect of the pre-arranged timetable persists and cannot be eliminated.

However, it is worth remembering that, nevertheless there are various not eliminable causes of delay (at least 16), it is possible to take them into account by building for each category of train the temporal distribution of presence in a determined point [3](Florio and Malavasi 1984), [4](Florio *et al.*, 1985).

Fig.1 shows an example of statistic distributions of arrival of trains in a singular point of a certain line and also the relative measures for the calculation of the average time distance. The figure shows the probability of train on time, while the distribution, indicated on the right, interprets the phenomenon of delay as a function of probability density. In these terms, each value of the interval $\bar{h}$ (Δt) has a probability to be overcome by an inferior interval, but also by a higher one. The value of $\bar{h}$ comes from considerations on the safety of circulation but a such value, if it can define the capacity of the line track system, could really be only theoretical, being often overcome in real management.

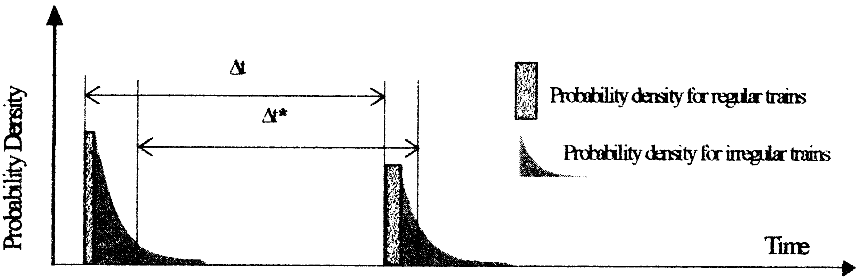

Figure 1: Probability density functions for regular and irregular trains.

3 Capacity of junctions and station tracks

As railway junction is intended a delimited area in a railway plant or system in which one train travelling along any paths prevents the circulation of one or more trains travelling on one or more paths incompatible with its own.

Observe the scheme in fig. 2 as an example. Each train movement from the left to the right or viceversa interdicts the circulation of other trains for an interval to be calculated according to the characteristics of the motion, if these are interested temporally by the interference. In other words, the railway junction works as a bottleneck in the railway flow as it allows an average number of movements (and, in certain cases, one at a time) that is always inferior that allowed by a system in parallel.

The first researcher on this subject [8](Potthoff, 1965) intended the junction, always uprooted from the railway system, as being interested only by an average number of movements to seek for every structural scheme through the use of the compatible paths matrix.

We wish to point out that every complex node must be split up into more elementary nodes [6](Florio and Mussone, 1995); for each one, the probabilities of interdiction must be calculated for each single train,

Transactions on the Built Environment vol 34, © 1998 WIT Press, www.witpress.com, ISSN 1743-3509

the relative capacity and the temporal use. Therefore, according to this formulation, in any complex node it could happen that only some movements reach saturation in T, while the others are still possible.

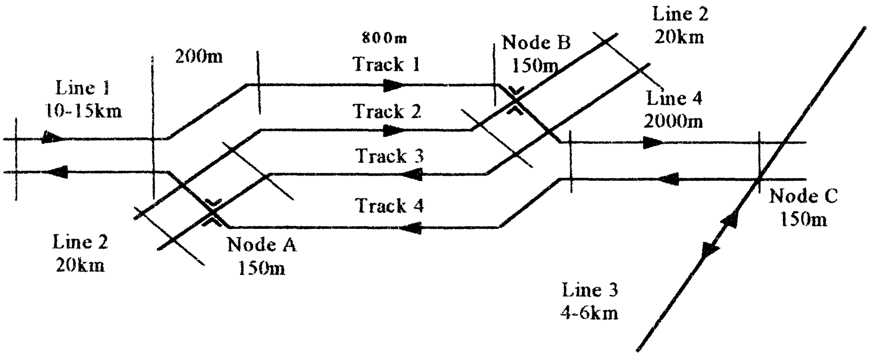

Figure 2: The railway scheme used for the example.

Once the splitting up of the complex junction in simpler nodes is made, even if they are mutually linked, it is relatively easy to calculate their dimension according to [5](Florio, 1992):

$$L_g = \frac{T}{\bar{t}_g} \qquad 2)$$

where L_g is the capacity of the node, $\bar{t}_g$ represents the average time of interdiction between trains. It is therefore necessary, once calculated $\bar{t}_g$ average time of interdiction between trains, to pass to the determination of $\bar{h}$, average interval between trains in each point marking the beginning or end of paths.

Such problem has been faced and solved in previous studies [6](Florio and Mussone, 1995) through the analytical splitting up of complex nodes in elementary ones and the solution of a problem of optimum whose aim is to maximise, in iterative form, the overall number of trains in the node, once the following data are known: all occupation times, probabilities of interference, capacity constraints on line and station tracks, constraints of congruence and not negativness of variables. It has been demonstrated that the problem of optimum is solved with further imposed limitations either on the number of trains on certain paths or on the ratio between them and/or on the categories they belong to.

In a generic station track -i - the occupation of every train is given temporally by a time t^i_{occr} function of the category r to which the train belongs. To be more precise, t^i_{occr} is made by three parts:

a) the real occupation time of the track, $t^i_{O,r}$, which includes the entrance-exit times and those needed to allow the passengers to get on/ get out and the usual operations in the station;
b) the additional time of delay caused by the line due to one of the (16 or more) possible causes of deviation from the timetable, $t^i_{R,r}$;
c) a further (possible) additional delay caused by the interference in the node in entry $t^i_{I,r}$.

In these terms, therefore $t^i_{occ,r}$ is given by the sum of three terms:

$$t^i_{occ,r} = t^i_{O,r} + t^i_{R,r} + t^i_{I,r} \qquad 3)$$

and the overall occupation of the track i in T is obtained according to:

$$\sum_{T,i} t^i_{occ,r} \qquad 4)$$

For each track, the following condition must be checked:

$$\sum_{T,i} t^i_{occ,r} \leq T \qquad 5).$$

4 The traffic vector of a railway system

From what is stated above, it can be assumed that, there is not only one value of capacity but a dominion of capacity for each component of railway scheme (line or station track, simple node part of a complex one).

The only condition that must be always verified is that, if U_e is the temporal use of every part in T, conditions of the following type are always verified:

$$U_e \leq T \qquad 6)$$

using the sum of the occupation times as in 5).

Actually, for every starting or ending point of the movements of trains also as any track stretch between two of them their number, divided into categories, can be expressed as a vector. The term vector gives the idea of stress of the scheme component and it is similar to the strain in Science of materials. In the same way, the capacity is the maximum effort or maximum solicitation admissible. It must be reminded that:
a) for every line track, the traffic vector has the number of trains in T compatible with the track capacity. It will display as many components as the possible categories of trains;
b) for each station track there will be a vector with as many components as the categories of trains, dimensionally similar to the previous one;
c) in each node, conditions of congruence must be present in such a way that the sum of the vectors of the confluent tracks is always equal to the

representative traffic vector on the track of convergence.

If, for instance, the vectors of the confluent tracks are called $A=[a_i]$ and $B=[b_j]$, for the vector C we will have:

$$C=[A+B]=[a_i+b_j] \qquad 7)$$

A similar procedure will be applied in the case of two or more tracks and to the difference between vectors in the case of tracks diverging in two or more.

5 Principles to work out the optimal time-table

When the traffic vectors compatible with the capacity of one or more components of the railway scheme are defined, it is necessary to draw up the "optimal" timetable, that is the one causing the minimum interference and the highest regularity of circulation or, at least, one that would not cause further delays and/or would contribute to cancel them.

Keeping in mind that the delays on line are, so to say, physiological and therefore impossible to cancel, it is possible to proceed according to two different criteria:

a) optimisation in the use of junctions. It is set that between two following traces of trains occupying the node there is always the same interval of time, δ, so that the time of not occupation of the node in T is fairly distributed;

b) optimisation of the service. It is set that the instant of arrival of the trains in the station obeys to the so-called repetitive mnemonic timetable (trains distributed in a continuous way during the whole interval, T, with a rigid period).

The first criterion tends to avoid that the node acts as a propagator of delays, while the second realises a quality optimum for the passenger, tending to the arrival of trains in station with a certain period. But, the second criterion implies the possibility of additional delays to those of line, as it can cause interference in the circulation in the entry nodes.

6 An application

The scheme adopted for the sake of example is in fig. 2. Three categories of trains are considered: fast (150m long), local (300m long) and freight (600m long). Stop time on station tracks is 2, 4 respectively for fast and local trains (freight trains don't stop), queue time due to irregularities are 8,5 and 3 m respectively for fast, local and freight trains. Speed used by no stopping trains on station tracks is 60km/h for the sake of this example.

The sample interval, T, is assumed equal to 240m. In the fig. 2 it can be recognised four lines the characteristics of which are:
line 1 t_{fm}=7m and t_{zu}=1 with a capacity of 30 trains each T period;
line 2 t_{fm}=5m and t_{zu}=1 with a capacity of 40 trains each T period;
line 3 t_{fm}=9m and t_{zu}=3 with a capacity of 10 trains each T period for each way;
line 4 t_{fm}=5m and t_{zu}=1 with a capacity of 40 trains each T period.

No fast train runs on line 3. On track 1 and 4 all trains stop; on tracks 2 and 3 only the 50% of trains stops (freight trains don't stop). In the same scheme three nodes, A, B and C, can be recognised. On the basis of these considerations three examples (scenarios) of management of the railway scheme are proposed and summarised in table 1.
The values of traffic capacity (vectors) are those respecting the capacity constraints on lines, junctions and stations. The first component to reach saturation in the three scenarios, and that limits the use of the others, is highlighted in table 1 and this is the real bottleneck of the system in the above-mentioned hypotheses.

Maximum Capacity Possible Vectors

Type	Fast	Local	Freight	Occupation %
Line 1	19	10	1	100
	21	7	2	100
	25	2	4	100
Line 2	27	13	0	100
&Line 4	30	8	2	100
	34	3	4	100
Line 3	0	16	4	100

First scenario

Type	Fast	Local	Freight	Occupation %
Line 1	9	5	1	54
Line 2	6	2	1	27
Line 3	0	4	2	33
Line 4	9	5	1	44
Track 1	9	5	1	70
Track 2	6	2	1	37
Track 3	6	2	1	37
Track 4	9	5	1	70
Node A	-	-	-	***100***
Node B	-	-	-	***100***
Node C	-	-	-	***100***

Second scenario

Type	Fast	Local	Freight	Occupation %
Line 1	89	4	1	47
Line 2	5	4	2	41
Line 3	0	8	2	50
Line 4	8	4	1	38
Track 1	8	4	1	60
Track 2	5	4	2	43
Track 3	5	4	2	43
Track 4	8	4	1	60
Node A	-	-	-	***100***
Node B	-	-	-	***100***
Node C	-	-	-	***100***

Third scenario

Type	Fast	Local	Freight	Occupation %
Line 1	2	2	0	14
Line 2	9	4	2	48
Line 3	***0***	***16***	***4***	***100***
Line 4	2	2	0	12
Track 1	2	2	0	19
Track 2	9	4	2	61
Track 3	9	4	2	61
Track 4	2	2	0	19
Node A	-	-	-	***100***
Node B	-	-	-	***100***
Node C	-	-	-	58

Table 1: Matrices of maximum capacity and different network scenarios.

7 Final Remarks

In this paper we have shown how all railway circulation characteristics can be treated analytically and in such a way that the set of railway components is considered as a system. Therefore lines and plants capacity can be treated according to the more general system theory. Moreover, we underline that capacity of each railway track can be seen as a vector to which usual algebraic rules can be applied.

In any case, our approach can be used to solve each railway problem. It can be demonstrated that as this procedure calculates the capacity and optimal time-tables for a given railway scheme, by reversing it, it can calculate the optimal railway scheme for a given demand (number of trains on line tracks).

Because the railway transport system is assimilable to other transport systems, it is easy to understand that the above-mentioned analytical formulas can be easily applied to them, with a few or no changes.

References

[1] Deutsche Bundesbahn (1974) Richtlinien fur die Ermittlung der Leistungsfahigkeit von Strecken(in German).

[2] Deutsche Bundesbahn (1979) Richtlinien fur die Ermittlung der Leistungsfahigkeit von Fahrstrassenknoten (in German).

[3] Florio L., Malavasi G. (1984) Principi teorici per la verifica di un impianto ferroviario complesso e la determinazione dei margini di potenzialità (in Italian). *Ingegneria Ferroviaria*, **12**, 846-852.

[4] Florio L., Malavasi G., Salvini E. (1985) Analisi dei ritardi e verifica di un nodo ferroviario complesso (in Italian). *Ingegneria Ferroviaria* ,**7**, 394-403.

[5] Florio L. (1992) La determinazione della potenzialità dei nodi ferroviari (in Italian). In L.Bianco, A.La Bella, eds., *Strumenti analitici per l'analisi dei sistemi di trasporto*, F.Angeli, Milano, pp. 375-407.

[6] Florio L., Mussone L. (1996) A method of capacity computation for complex railway systems. 7th WCTR, *World Conference on Transport Research*, Vol 4, pp. 275-292, Elsevier Science, 1996.

[7] Petersen E.R., Taylor A.J.(1982) A Structured Model for Rail Line Simulation and Optimization. *Transportation Science* ,**2**, 192-206.

[8] Potthoff G.(1965) Verkehrsstromungslehre (in German). Voll.1-5 Transveb, Berlin.

[9] U.I.C. (1978) *Fiche 405R.*