

Baseline RailNL

Menno Rooker, Ron Lakeman & Jero Tretmans
”Spoorjewel”

January 2024

Introductie

Voor de baseline van onze case is een random algoritme geïmplementeerd om een idee te krijgen van de gemiddelde score van een dienstregeling die voor RailNL gemaakt kan worden. De scores zijn bepaald volgens de doelfunctie:

$$K = p * 10000 - (T * 100 + Min) \quad (1)$$

waarin K de kwaliteit van de lijnvoering is, p de fractie van de bereden verbindingen (dus tussen 0 en 1), T het aantal trajecten en Min het aantal minuten in alle trajecten samen.

Het Algoritme

Het random algoritme dat is geïmplementeerd kiest eerst een willekeurig aantal (tussen 1 en 7) trajecten dat het gaat maken. Vervolgens kijkt het naar de beschikbare connecties en maakt daar een willekeurige keuze uit. Ten slotte moet het algoritme bepalen wanneer het een traject stopt. Dit doet het algoritme door telkens een bovengrens aan het aantal minuten per traject te variëren met een `randint` tussen 1 en 120. Met deze constraints is het uiteindelijke resultaat is een ”semi-willekeurige” dienstregeling met een bijbehorende score. Veel van zulke scores bij elkaar geven een idee van wat voor scores een geoptimaliseerd algoritme zou moeten overtreffen.

Visualisatie

Hieronder is een visualisatie van de scores die het random algoritme heeft verworven. Het is duidelijk dat de histogram right-skewed is, wat betekent dat de waarden vaak relatief laag liggen. Dit is op zichzelf niet erg aangezien dit algoritme puur bedoelt is om een ondergrens vast te stellen voor de andere algoritmes die geïmplementeerd gaan worden. Een groot aantal scores zit onder de 2000 maar met een gefitte lijn krijgen we een idee van de gemiddelde score, deze lijkt tussen de 2000 en 3000 te liggen.

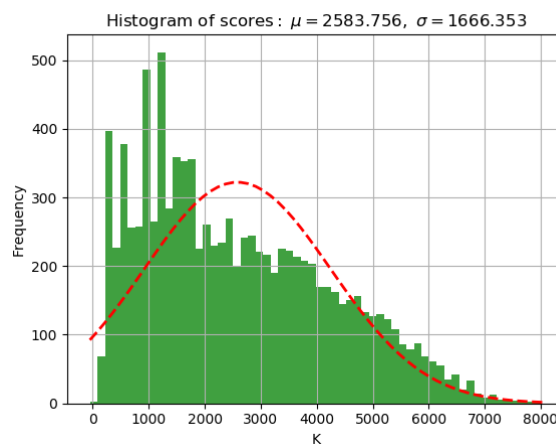


Figure 1: Scorefrequenties van het random algoritme

Reflectie

Intuïtief gezien zou je verwachten dat een volledig willekeurig algoritme ook een volledig willekeurige grafiek zou produceren. Natuurlijk is het zo dat er een aantal constraints moeten worden gesteld aan de case, waardoor volledige willekeur niet mogelijk is, althans niet voor het verwerven van daadwerkelijke states die oplossingen van de case zouden kunnen zijn. Om ervoor te zorgen dat de resultaten van het algoritme nuttig zijn hebben we al wat beperkingen toegepast, zoals in de opdracht omschreven staat. Die beperkingen bevatten het limiteren van het aantal trajecten tot een maximum van 7, en ervoor zorgen dat elk traject tussen de 1 en 120 minuten is. Er is geen voorkeur gegeven aan langere trajecten of meer trajecten, ook kan een trein continu op en neer reizen op hetzelfde traject waardoor de tijd oploopt zonder extra unieke verbindingen te bereiken. Deze factoren geven een verklaring voor de afwijkende data.