

Voorspelling netwerkverkeer

met behulp van

klantgroepen en tijdreeksen



Master project Business Analytics

ing. R.C.Lam



Voorspelling netwerkverkeer

met behulp van

klantgroepen en tijdreeksen

Master project Business Analytics

ing. R.C.Lam

KPN B.V.
Koninklijke PTT Nederland
Stationsstraat 115
3811MH Amersfoort

Vrije Universiteit Amsterdam
Faculteit der Exacte Wetenschappen
De Boelelaan 1105
1081 HV Amsterdam

Begeleiders:
P. Koeleman
M. de Nooijer
R. van der Mei
F. Bijma

november 2015



Voorwoord

Voor u ligt mijn scriptie die is geschreven ter afsluiting van de master Business Analytics aan de Vrije Universiteit Amsterdam. Hierin zijn de resultaten van mijn afstudeerstage beschreven die ik heb uitgevoerd bij KPN, de Koninklijke PTT Nederland, van juni tot en met november 2015.

Er zijn meerdere personen die ik wil bedanken. Allereerst mijn begeleidster Paulien Koeleman van de KPN. Zij heeft wekelijks mijn voortgang bijgehouden en kwam met positieve feedback. Hierdoor kreeg ik weer nieuwe input en kon ik verder met mijn onderzoek. Paulien is erg resultaatgericht en dit helpt tot mooie inzichten te komen.

Na drie maanden kreeg mijn stage een bijzondere wending. Paulien kreeg een andere baan aangeboden en heeft besloten KPN te verlaten. Marco de Nooijer is de manager van mijn afdeling, Business Intelligence, en heeft besloten mijn stage de laatste drie maanden te begeleiden. Ik wil hem bedanken voor de mogelijkheid om dit afstudeerproject te doen en de feedback op mijn scriptie en eindproduct. Uiteraard wil ik graag alle collega's van KPN bedanken voor de positieve werkomgeving.

Daarnaast wil ik ook mijn begeleiders van de Vrije Universiteit bedanken. Rob van der Mei voor de ondersteuning en feedback. Hij is een aantal keer bij de KPN langsgekomen en ik kon hem alles vragen wat in mij opkwam. Tot slot wil ik ook Fetsje Bijma bedanken, zij is de tweede lezer vanuit de Vrije Universiteit.

Ruben Lam

november 2015

Samenvatting

Het optimaal uitbreiden van het netwerk is cruciaal voor KPN. Enerzijds om alle klanten van internet en TV te voorzien en daarnaast om de kosten van uitbreidingen zo laag mogelijk te houden. Wanneer het netwerk onnodig uitgebreid wordt, is er sprake van overdimensionering. Dit zal KPN zeer hoge kosten opleveren aangezien jaarlijks 400 miljoen euro aan netwerkuitbreiding wordt uitgegeven.

De huidige methode voor uitbreiding gebeurt handmatig. Er wordt wekelijks een berekening gemaakt om te bepalen welke netwerkonderdelen op meer dan 75% van hun capaciteit zitten. Dit is tijdrovend en inefficiënt. Deze scriptie beschrijft een methode om dit proces te verbeteren.

In dit verslag komen verschillende algoritmes voor het voorspellen van tijdreeksen aan bod. Deze tijdreeksen bevatten informatie over het netwerkverbruik per locatie in Nederland. De algoritmes die beschreven zijn en gebruikt worden zijn het Holtwinters algoritme, de STL methode, ARIMA en ARIMA met Fourier.

Het resultaat is een dynamisch model dat op basis van een errormeting een algoritme kiest per tijdserie. De voorspelling die dit model maakt kan gebruikt worden voor optimale netwerkuitbreiding. Daarnaast geeft het inzicht in het verloop van het netwerkverkeer, zowel voor het verleden als de toekomst.

Door het gebruik van dit model kan een slimme uitbreidingsstrategie bepaald worden en kostenbesparing worden gerealiseerd.