

Stedelijke verkeerspatronen

Wendy Weijermars & Eric van Berkum

Vakgroep Verkeer, Vervoer en Ruimte

Universiteit Twente

w.a.m.weijermars@utwente.nl

Bijdrage aan het Colloquium Vervoersplanologisch Speurwerk 2004,

25 en 26 november 2004, Zeist

Inhoudsopgave

1. Inleiding.....	4
2. Stedelijke verkeerscentrales	4
<i>Data</i>	5
<i>Verkeerscentrale</i>	6
<i>Verkeersinformatie.....</i>	6
<i>ViaContent</i>	7
3. Analyse verkeerspatronen	8
<i>Definitie verkeerspatroon.....</i>	8
<i>Clusteranalyse.....</i>	10
<i>Beoordeling clusteringen</i>	11
<i>Analyse waargenomen verkeerspatronen.....</i>	12
<i>Voorspellen verkeerspatronen</i>	12
4. Voorbeeld clustering	12
<i>Omschrijving ontstane clusters</i>	14
<i>Beoordeling clusteringen</i>	14
<i>Nadere analyse ontstane clusters.....</i>	16
5. Conclusie	17
Referenties	19

Samenvatting

Stedelijke verkeerspatronen

Mede dankzij de ontwikkeling van stedelijke verkeerscentrales komt de laatste jaren steeds meer informatie over het verkeer op het stedelijke wegennet beschikbaar. De ingewonnen gegevens kunnen ondermeer gebruikt worden voor de analyse van verkeerspatronen die vervolgens gebruikt kunnen worden voor de ontwikkeling van verkeersmanagementscenario's en het voorspellen van intensiteiten. Het definiëren en analyseren van verkeerspatronen verdient echter wel de nodige aandacht. Dit paper beschrijft de analyse van dagelijkse intensiteitspatronen. Een patroon wordt daarbij in eerste instantie gedefinieerd als een serie kwartierintensiteiten. Daarnaast wordt het patroon ook gedefinieerd aan de hand van spitsintensiteiten, totale intensiteit, spitsstijden en verhoudingen tussen dagdelen. Door middel van clustering kunnen de patronen in groepen worden ingedeeld. Uit een verkennende analyse op een beperkte hoeveelheid data kan geconcludeerd worden dat de clustering op basis van uurintensiteiten leidt tot dagpatronen die duidelijk van elkaar verschillen qua vorm en hoogte. De variatie binnen de patronen blijkt echter aanzienlijk te zijn. Wellicht leidt een clustering op basis van bijvoorbeeld spitsintensiteiten, totale intensiteit en spitsstijden tot betere resultaten. Uit nadere analyse van de ontstane clusters blijkt ten slotte dat de bouwvakvakantie een bepalende factor is voor de clustering. Daarnaast blijken vrijdagen in een ander cluster te worden ingedeeld dan maandagen, dinsdagen en donderdagen. Deze informatie kan gebruikt worden om vooraf in te schatten hoe een dagpatroon er op een bepaalde dag uit zal zien.

Summary

Urban traffic patterns

Urban traffic information centers provide a lot of information about the traffic situation on the urban transport network. This information can be used to detect traffic patterns. These traffic patterns can be used for the development of traffic management scenario's and for traffic forecasts. This paper describes a method to analyze traffic patterns. First of all, traffic patterns are defined. Our focus is on daily traffic profiles that are defined in two ways. According to the first definition, a daily flow profile consists of a series of (quarter hourly) traffic flows. The second definition uses peak hour flows, daily traffic flow, time interval of peak periods and ratios between periods of the day as defining variables. Clustering analysis is used to distinguish different traffic patterns. In a preliminary analysis, a limited amount of days is clustered. When hourly traffic flows are used as features, the clustering leads to distinctive patterns concerning the height and shape of the patterns. The variation within the patterns is however substantial. Probably, a clustering based on peak flows, daily flows, peak times and ratios provides better results. The construction industry holiday appears to be responsible for one of the resulting clusters. Furthermore, Fridays are in general divided in a distinct cluster from Mondays, Tuesdays and Thursdays. This information can be used to estimate daily traffic patterns in advance.

1. Inleiding

De ontwikkeling van stedelijke verkeerscentrales heeft de laatste jaren in sneltreinvaart plaatsgevonden. In verschillende Europese projecten is aandacht besteed aan de ontwikkeling van stedelijke informatiesystemen (bijvoorbeeld SCOPE (Clarke & Morris, 2001), CAPITALS(+), ENTERPRICE (Kirschfink et al, 2000) and QUARTET PLUS). Daarnaast wordt in veel steden op eigen initiatief een informatiesysteem ontwikkeld.

Bijkomend voordeel van stedelijke verkeerscentrales is dat ze een schat aan informatie over de verkeerssituatie leveren. In steeds meer steden wordt deze verkeersinformatie gebruikt voor verkeersmanagement. Verkeersgegevens die over een langere termijn verzameld zijn, worden gebruikt om inzicht te verkrijgen in de aanwezige verkeerspatronen. Voor deze verkeerspatronen kunnen vervolgens verkeersmanagementscenario's ontwikkeld worden. Ook kan het inzicht in de aanwezige verkeerspatronen gebruikt worden voor voorspellingen. Deze voorspellingen kunnen vervolgens weer gebruikt worden voor reisinformatie.

Het definiëren en analyseren van stedelijke verkeerspatronen verdient echter wel de nodige aandacht. Allereerst moet bepaald worden wat onder een verkeerspatroon verstaan wordt. Vervolgens moeten verschillende patronen onderscheiden worden en moet onderzocht worden welke factoren ten grondslag liggen aan de gevonden patronen.

Dit paper beschrijft hoe de door een verkeerscentrale geproduceerde gegevens gebruikt kunnen worden voor de analyse van stedelijke verkeerspatronen. Het eerste hoofdstuk gaat in op stedelijke verkeerscentrales. In het tweede hoofdstuk wordt vervolgens aangegeven hoe de verkeerspatronen geanalyseerd worden. In het derde hoofdstuk wordt een voorbeeld gegeven van een uitgevoerde analyse. Het paper wordt afgesloten met een conclusie.

2. Stedelijke verkeerscentrales

Het basisprincipe van een verkeerscentrale is als volgt:

Data → verkeerscentrale → verkeersinformatie

Deze onderdelen worden in de volgende paragrafen behandeld aan de hand van gevonden informatie over bestaande stedelijke verkeerscentrales. Tot slot wordt kort ingegaan op ViaContent, het systeem dat de verkeersgegevens levert voor de analyses van de verkeerspatronen.

Data

De verzamelde gegevens verschillen van stad tot stad. In alle steden worden in ieder geval verkeersgegevens verzameld. Dit kunnen gegevens zijn ten aanzien van aantallen voertuigen, bezettingsgraden, snelheden en verkeerslichtenregelingen. Voor de inwinning van verkeersgegevens zijn verschillende systemen beschikbaar. Ten eerste wordt in de meeste steden detectordata afkomstig van inductielussen, infrarood- of radardetectors gebruikt (bijvoorbeeld Budde, 2002; Richards et al, 2001; Scharrer et al, 2003). Daarnaast worden ook regelmatig data afkomstig van (CCTV) camera's (bijvoorbeeld Ancidei et al, 2000; Cone et al, 2002; Karl & Trayford, 2000), probe vehicles (bijvoorbeeld Bae & Lee, 2000; Fellendorf et al, 2000; Ferulano et al, 2000) en GPS of GSM (Karl & Trayford, 2000; Leitsch, 2002) gebruikt. Tot slot zijn er enkele steden met minder gangbare informatiebronnen gebruikt zoals verkeersdeelnemers die informatie doorgeven (Bae & Lee, 2000), openbaar vervoer of systemen die de prioriteit van het openbaar vervoer regelen (Henriet & Schmitz, 2000; Hoyer & Herrmann, 2003) en helicopters (FHWA, 2003).

Naast verkeersgegevens worden ook vaak andere gegevens verzameld. Vaak wordt informatie toegevoegd ten aanzien van wegwerkzaamheden, evenementen en incidenten (bijvoorbeeld Hasberg & Serwill, 2000; Leitsch, 2002) zodat reizigers een volledig beeld krijgen van de toestand op het wegennet. Daarnaast wordt een stedelijk verkeersinformatiesysteem vaak gekoppeld aan een parkeerinformatiesysteem (bijvoorbeeld Budde, 2002). Ook zijn er multimodale verkeerscentrales (bijvoorbeeld Hasberg & Serwill, 2000) die informatie verschaffen over verschillende modaliteiten en zo trachten een modal shift teweeg te brengen. In sommige gevallen worden daarnaast ook weergegevens (Cone et al, 2002; Kellermann & Schmid, 2000), gegevens met betrekking tot het milieu (Ancidei et al, 2000; Kellermann & Schmid, 2000) of een kalender (Kellermann & Schmid, 2000) bijgehouden.

Verkeerscentrale

In de verkeerscentrale worden verschillende databewerkingen uitgevoerd alvorens verkeersinformatie gegenereerd wordt. Sommige centrales beperken zich tot het controleren en samenvoegen van data uit verschillende bronnen. De controlealgoritmen variëren van simpele procedures die onrealistische metingen verwijderen tot meer geavanceerde systemen waarbij de verzamelde gegevens gecontroleerd worden aan de hand van historische data of data van meerdere meetpunten. In sommige steden (bijvoorbeeld Bae & Lee, 2000; Henriët & Schmitz, 2000) wordt data uit verschillende bronnen met elkaar vergeleken om zo tot een zo betrouwbaar mogelijk databestand te komen. In Barcelona (FHWA, 2003) worden reistijden geverifieerd door een “floating car”. In veel verkeerscentrales worden de gegevens daarnaast gevisualiseerd.

Meer geavanceerde centrales gebruiken verkeersmodellen om de verkeerssituatie op het gehele wegennet te bepalen of om reistijden te schatten of te voorspellen. In München bijvoorbeeld, is de MOBINET Control and Information Centre (MCIC) geïnstalleerd. Dit systeem bepaalt op basis van verkeerstellingen en historische Herkomst-Bestemmingsinformatie de actuele verkeerssituatie in de vorm van een HB-matrix, reistijden of level of service (Kellerman and Schmid, 2000). In Turijn schat de Town Supervisor (MIZAR Automazione, 1998) dynamische HB-matrices en intensiteiten in het gehele netwerk op basis van verkeersgegevens die door verschillende subsystemen gemeten worden. In Hannover voorspelt VISUM-online op basis van alle beschikbare historische en recente verkeersgegevens de huidige en toekomstige verkeerssituatie (Fellendorf et al, 2000). VISUM-online maakt hiervoor gebruik van algoritmen die dynamische HB matrices schatten op basis van verkeerstellingen. In Keulen wordt de data geproduceerd door MOTION -het systeem dat de verkeersregelingen optimaliseert- en wordt het incidentdetectiealgoritme van MOTION gebruikt om congestie op te sporen en reizigers hierover te informeren (Kruse et al, 2000).

Verkeersinformatie

De verkeerscentrales in de verschillende steden kennen diverse toepassingen. Een functie die in alle gevallen terugkomt, is het verschaffen van reisinformatie aan de reiziger. Hiervoor

worden verschillende media gebruikt zoals Internet, radio, mobiele telefoon, on-board navigatiesystemen en dynamische routeinformatiepanelen.

Naast het verstrekken van verkeersinformatie aan reizigers, kan de verkeerscentrale ook gebruikt worden voor verkeersmanagement, verkeersplanning (Leitsch, 2002), beleidsondersteuning en onderzoek. Met betrekking tot verkeersmanagement kan onderscheid gemaakt worden in online in offline management. Bij online verkeersmanagement wordt gebruik gemaakt van actuele verkeersgegevens om direct in te grijpen in het verkeerssysteem door bijvoorbeeld verkeersregelingen te optimaliseren (bijvoorbeeld Ferulano et al, 2000; MIZAR Automazione, 1998). Offline verkeersmanagement omhelst het ontwikkelingen van maatregelen en regel- en informatiestrategieën met behulp van verkeersgegevens die over een langere periode verzameld zijn. In Wales zijn in het kader van pro-actief verkeersmanagement bijvoorbeeld verschillende informatiestrategieën ontwikkeld (Cone et al, 2002).

In een geavanceerde verkeersmanagementcentrale worden offline en online verkeersmanagement gecombineerd om tot een zo optimaal mogelijke verkeersafwikkeling te komen. Met behulp van historische gegevens kan een aantal verkeersmanagementscenario's ontwikkeld worden. Actuele verkeersgegevens kunnen vervolgens gebruikt worden om verder te optimaliseren rekening houdend met de actuele situatie. MacLennan & Routledge (1995) presenteren een voorstel voor een dergelijk systeem. In het Urban Traffic Management and Control (UTMC) System dat zij beschrijven wordt data van verschillende bronnen gecombineerd en wordt de verkeerssituatie voorspeld. Afhankelijk van de geldende verkeerssituatie kan vervolgens een van de –vooraf gedefinieerde- regelstrategieën geselecteerd worden. Ook in Europese projecten zijn verschillende geavanceerde verkeersmanagementcentrales ontwikkeld. Deze systemen bevatten naast het verzamelen en bewerken van verkeersinformatie ook onderdelen voor intelligente data-analyse en beslissingondersteuning (Kirschfink et al (2000); Barcelo (2000)).

ViaContent

Viacontent verzamelt, interpreteert en presenteert verkeersinformatie. De verkeersregelinstallaties vormen de belangrijkste bron voor informatie. Daarnaast kunnen ook andere databronnen ingezet worden, zoals telsystemen in parkeergarages en Floating Car

Data. Met behulp van programma's die op basis van verkeersmodellen werken, wordt de data bewerkt. De verkregen verkeersinformatie kan gebruikt worden door wegbeheerders in het kader van verkeersmanagement en kan worden doorgespeeld aan reizigers via bijvoorbeeld internet en routepanelen. Meer informatie is beschikbaar op www.viacontent.nl.

Door ViaContent gemeten verkeersintensiteiten in Almelo worden gebruikt voor de analyse van stedelijke verkeerspatronen. Er worden kwartierintensiteiten¹ per detector aangeleverd. Deze intensiteiten worden gecontroleerd en geaggregeerd naar linkniveau. De intensiteiten worden vervolgens gekoppeld aan kalendergegevens, wegwerkzaamheden, evenementen, ongevallen en weergegevens. Gegevens met betrekking tot evenementen en wegwerkzaamheden worden geleverd door de gemeente Almelo, gegevens met betrekking tot ongevallen worden geleverd door de politie en de weergegevens zijn afkomstig van het KNMI.

3. Analyse verkeerspatronen

Alvorens verschillende patronen onderscheiden kunnen worden, moet een patroon eerst gedefinieerd worden. Deze definitie wordt behandeld in de eerste paragraaf van dit hoofdstuk. De tweede paragraaf behandelt de daadwerkelijke indeling in groepen. De derde paragraaf behandelt de beoordeling van de ontstane clusters. Voor de clustering wordt gebruik gemaakt van historische data. Wanneer de verkeerspatronen echter gebruikt worden voor verkeersmanagement, is het van belang vooraf te weten welk patroon zich op een bepaald moment voordoet. In de derde paragraaf wordt daarom beschreven hoe de verkeerspatronen geanalyseerd kunnen worden. De vierde paragraaf gaat ten slotte in op het gebruik van verkeerspatronen voor verkeersvoorspellingen.

Definitie verkeerspatroon

Afhankelijk van de toepassing, kunnen verschillende soorten patronen onderscheiden worden. Allereerst kan onderscheid gemaakt worden in temporele en spatiele patronen. Dit paper gaat dieper in op temporele verkeerspatronen.

Temporele verkeerspatronen beschrijven de ontwikkeling van intensiteiten in de tijd. Daarbij kunnen verschillende tijdseenheden worden beschouwd. Gekeken kan bijvoorbeeld worden naar de variatie van minuut tot minuut of naar de variatie van jaar tot jaar. Wij zijn geïnteresseerd in de variatie van dag tot dag. Daarbij gaat het ons in de eerste plaats om de ontwikkeling van de intensiteiten gedurende de dag. Naast het dagpatroon kan het ook interessant zijn om de spitsen op verschillende dagen met elkaar te vergelijken. Daarom worden ook het ochtend- en avondspitspatroon beschouwd.

Een dagpatroon kan in de eerste plaats gedefinieerd worden als een serie intensiteiten als functie van de tijd. Het aantal intensiteiten is afhankelijk van het gewenste aggregatieniveau. Wanneer voor een laag aggregatieniveau gekozen wordt, worden zowel korte-termijn als lange-termijn variaties meegenomen, terwijl bij een hoger aggregatieniveau de korte-termijn variaties uitgefilterd worden. De korte-termijn variaties worden in de stad veroorzaakt door kleine verstoringen zoals verkeerslichten of een lossende vrachtwagen. Lange-termijnvariaties zijn voornamelijk het gevolg van variatie in verkeersvraag. Omdat wij geïnteresseerd zijn in variaties als gevolg van variaties in verkeersvraag, wordt gekozen voor een aggregatieniveau van 15 minuten.

Nadeel van bovenstaande definitie van een dagpatroon is dat deze weinig inzicht verschaft in het type verschillen tussen patronen. Daarom hebben wij een dagpatroon ook gedefinieerd aan de hand van een aantal kenmerken waarop dagen van elkaar kunnen verschillen. Allereerst zijn voor verkeersmanagementdoeleinden de piekintensiteiten erg belangrijk. Deze worden daarom ook beschouwd en bepalen gezamenlijk met de dagintensiteit de hoogte van het dagpatroon. Naast de hoogte speelt ook de vorm van het patroon een rol. Deze factor wordt meegenomen middels de spitsstijden en de verhoudingen tussen verschillende dagdelen. Tot slot moet nog gedefinieerd worden wat onder een spits verstaan wordt. Wij maken onderscheid in een absolute spitsperiode die is gedefinieerd als de twee opeenvolgende drukste uren en de relatieve spitsperiode die is gedefinieerd als de periode waarin de intensiteit hoger is dan twee maal de gemiddelde intensiteit.

¹ De intensiteiten worden iedere 5 minuten opgeslagen, maar na overleg is besloten dat voor dit onderzoek gebruik gemaakt wordt van kwartierintensiteiten.

Voor de definitie van spitspatronen wordt gebruik gemaakt van een derde definitie van spitsperiode, namelijk de drie opeenvolgende drukste uren. Op deze wijze kunnen ook de opbouw en de afbouw van de spits geanalyseerd worden. Een spitspatroon wordt vervolgens gedefinieerd als de opeenvolgende intensiteitsmetingen gedurende de periode en als de gemiddelde spitsintensiteit, de piekintensiteit en het tijdsinterval met de hoogste intensiteit.

Clusteranalyse

De patronen worden met behulp van clusteranalyse ingedeeld in groepen. De uitkomst van de clustering wordt sterk beïnvloed door de variabelen die het patroon vastleggen en de methode die gebruikt is voor de clustering.

De variabelen die het patroon vastleggen worden bepaald door de patroondefinitie. Wanneer wordt uitgegaan van de eerste patroondefinitie wordt het patroon vastgelegd door de opeenvolgende intensiteiten. Wanneer wordt uitgegaan van de tweede patroondefinitie wordt het patroon vastgelegd door de totale intensiteit, de spitsintensiteiten, de spitstijden en de verhoudingen tussen de verschillende dagdelen. Probleem hierbij is echter dat deze variabelen niet zomaar naast elkaar gebruikt kunnen worden. Ze worden namelijk niet op eenzelfde schaal gemeten en moeten daarom eerst gestandaardiseerd worden. Daarbij moet ook bepaald worden hoe zwaar een verschil in spitstijd weegt ten opzichte van een bepaald verschil in intensiteit. Omdat dit een zeer moeilijke afweging is, hebben wij gekozen voor een getrapte clustering waarbij patronen eerst geclusterd worden op basis van bijvoorbeeld de gemiddelde intensiteit en de ontstane clusters vervolgens verder onderverdeeld worden op basis van de spitstijden. Bijkomend voordeel van een dergelijke clustering is dat het inzicht in type verschillen tussen de clusters behouden blijft.

In Jain et al (1999) worden verschillende methoden voor clustering besproken. Allereerst kan onderscheid gemaakt worden in hiërarchische en partitionele methoden. Bij hiërarchische clusteringen worden verschillende indelingen geproduceerd, waaruit achteraf gekozen wordt. Partitionele clusteringen leiden tot één enkele indeling. Nadeel van deze methode is dat het aantal clusters vooraf gekozen moet worden. Om deze reden hebben wij gekozen voor een hiërarchische clustering. Het optimale aantal clusters wordt vervolgens bepaald met behulp van het dendrogram. Hierin wordt per clustering aangegeven wat de gemiddelde afstand

binnen de clusters is. Er wordt gekozen voor de clustering waarbij een verdere afname van het aantal clusters leidt tot een sterke toename van de gemiddelde afstand tussen patronen binnen de nieuwe clusters.

In de literatuur worden de volgende hiërarchische clustering procedures onderscheiden (Webb, 2002; Manchester Metropolitan University):

- Average linkage clustering; het verschil tussen clusters wordt bepaald door de gemiddelde waarden
- Complete linkage clustering; het verschil tussen clusters wordt bepaald door het grootste verschil tussen leden van beide clusters
- Single linkage clustering; het verschil tussen clusters is gelijk aan de minimale afstand tussen leden van beide clusters
- Within groups clustering; bij deze methode wordt variatie binnen een cluster geminimaliseerd
- Ward's method; hierbij is de clustering gebaseerd op de som van de kwadratische afwijkingen ten opzichte van het midden van het cluster.

Van bovenstaande methoden is de single linkage clustering duidelijk minder geschikt in dit geval. Bij deze methode kunnen patronen die niet op elkaar lijken namelijk in hetzelfde cluster worden ingedeeld omdat er toevallig tussenliggende patronen zijn die de minimale afstand tussen clusters verkleinen. De overige methoden zijn onzes inziens allemaal geschikt. Wij hebben gekozen voor Ward's method omdat deze vaak gebruikt wordt en in het verleden tot goede resultaten gekomen is [Nowotny et al, 2003].

Beoordeling clusteringen

De clusteringen worden tot slot beoordeeld aan de hand van de variatie in dagpatronen binnen de clusters. Wanneer de ontstane groepen gebruikt worden voor de ontwikkeling van verkeersmanagementscenario's is het namelijk belangrijk dat de dagen binnen een groep een soortgelijk dagpatroon vertonen en ook voor verkeersvoorspellingen is het belangrijk dat de patronen binnen een groep zoveel mogelijk op elkaar lijken.

Analyse waargenomen verkeerspatronen

Zoals al in de inleiding naar voren is gekomen, kan met behulp van bovenstaande clustering pas achteraf bepaald worden tot welke groep een bepaald dagpatroon behoort. Voor een effectieve inzet van verkeersmanagementscenario's is het van belang te weten wanneer bepaalde verkeerspatronen zich voordoen. Daarom wordt onderzocht welke factoren ten grondslag liggen aan de ontstane clusters.

Verwacht wordt dat de volgende kenmerken het waargenomen intensiteitenpatroon kunnen beïnvloeden:

- Kenmerken met betrekking tot het type dag; dag van de week, seizoen, feestdag, vakantieperiode
- Kenmerken met betrekking tot specifieke omstandigheden; wegwerkzaamheden, ongevallen, evenementen
- Kenmerken met betrekking tot het weer; hoeveelheid en duur van de neerslag, mate bewolking en zon, temperatuur

Voor deze kenmerken wordt dan ook onderzocht in hoeverre ze verantwoordelijk zijn voor de ontstane clusters.

Voorspellen verkeerspatronen

Bij pattern-based forecasting wordt de verkeersintensiteit voorspeld met behulp van patroonherkenning. Het waargenomen intensiteitsverloop worden ingedeeld in één van de patronen die in een eerdere stap gedefinieerd zijn. Om het herkennen van patronen te vergemakkelijken kan gebruik gemaakt worden van preclassificatie. De factoren die in de vorige paragraaf behandeld zijn kunnen in dat geval gebruikt worden om een eerste indeling in groepen te maken. De waargenomen intensiteiten worden vervolgens gebruikt voor een verder onderscheid in meer specifieke patronen. Wild (1997) classificeert de dagen eerst op basis van de dag van de week en evenementen die plaatsvinden, alvorens de dagen door middel van clustering verder in te delen in specifiekere groepen.

4. Voorbeeld clustering

Op dit moment is de data uit ViaContent nog niet beschikbaar. Wel is er een testset beschikbaar met intensiteitgegevens van een aantal kruispunten in Almelo. Deze gegevens

zijn verzameld in de periode augustus – oktober 2002. Op deze testset zijn analyses uitgevoerd om het inzicht in stedelijke verkeersstromen te vergroten. Een uitgebreidere beschrijving van de data en de uitgevoerde verkennende analyse is te vinden in Weijermars & Van Berkum (2004).

Op één van de bestudeerde verkeersstromen is een clusteranalyse uitgevoerd. Het gaat om de verkeersstroom op de H. Roland Holstlaan richting de autosnelweg (A35) (zie figuur 1).



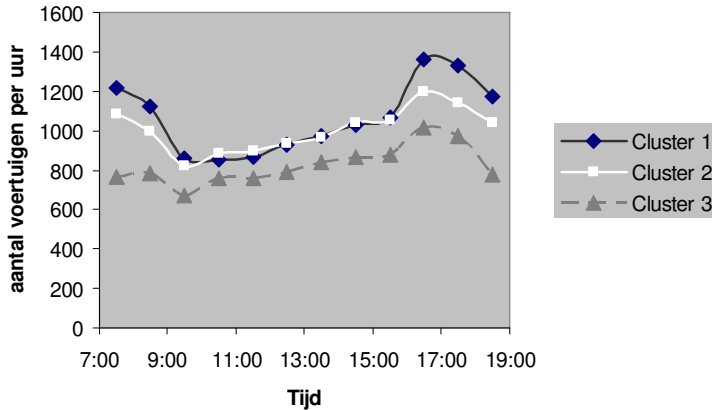
Figuur 1: Geanalyseerde verkeersstroom (bron: Michelin2004 © Teleatlas)

Opgemerkt moet echter wel worden dat de clustering is uitgevoerd op een beperkt aantal dagen. Er is data verzameld op 56 dagen, waarvan 27 dagen binnen de vakantieperiode gelegen zijn. Daarnaast zijn alleen uurintensiteiten beschikbaar en kan dus geen nauwkeurig dagpatroon geschetst worden. De resultaten dienen daarom slechts ter illustratie van de omschreven methode. Vanwege de beperkte hoeveelheid gegevens heeft alleen een clustering op basis van uurintensiteiten plaatsgevonden. Een getrapte clustering op basis van meerdere kenmerken leidt tot meer clusters en in dit geval bleken al snel te kleine clusters te ontstaan die niet verder kunnen worden onderverdeeld. Ook de spitsperioden zijn niet in clusters ingedeeld aangezien het aggregatieniveau te hoog is om de ontwikkeling van de spits goed te beschrijven.

Eerst wordt een omschrijving gegeven van de ontstane clusters en vervolgens wordt nader onderzocht welke kenmerken aan de grondslag van de ontstane clusters liggen.

Omschrijving ontstane clusters

Bestudering van het dendogram leidt tot een indeling in 3 clusters. De gemiddelde dagpatronen van deze clusters zijn weergegeven in figuur 2. Daarnaast geeft tabel 1 de gemiddelde dag- en piekintensiteiten voor de drie clusters.



Figuur 2: dagpatronen clusters

Cluster	Aantal dagen	Totale intensiteit	Ochtendspits	Avondspits
1	27	12795	1219	1375
2	16	12070	1091	1203
3	13	9877	826	1069

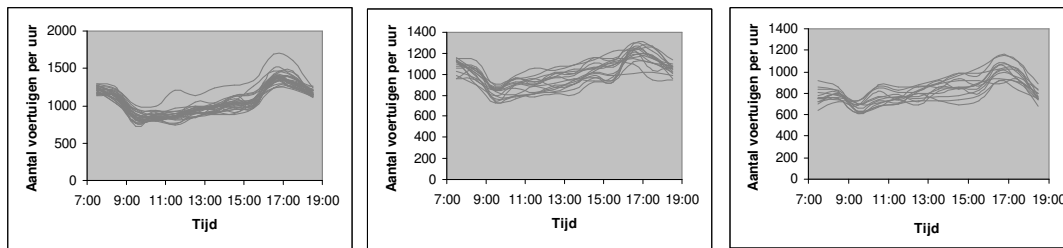
Tabel 1: Kenmerken clusters

Een gemiddelde dag in cluster 3 heeft gedurende de hele dag duidelijk een lagere intensiteit dan dagen in de andere clusters. De totale intensiteit is voor cluster 2 gemiddeld lager dan voor cluster 1. Uit de grafiek blijkt dat dit verschil veroorzaakt wordt tijdens de spitsperiodes. Het dagpatroon van cluster 2 vertoont een vlakker verloop dan van cluster 1. Cluster 3 vertoont een nog vlakker verloop dan de andere clusters. Wanneer de ochtendspits gedefinieerd wordt als periode dat de intensiteit hoger is dan tweemaal de gemiddelde intensiteit, blijkt dat bijna de helft van de dagen in cluster 3 geen ochtendspits kent. Ook de verhouding tussen de ochtendspitsintensiteit en de intensiteit tussen de spitsen is lager voor cluster 3 dan voor de andere clusters.

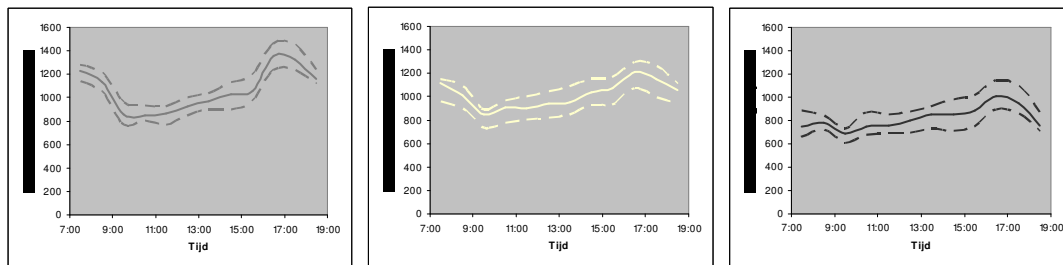
Beoordeling clusteringen

De clusteringen worden zoals –zoals vermeld- beoordeeld aan de hand van de variatie tussen dagen binnen de clusters. Allereerst zijn de dagen binnen de afzonderlijke clusters

weergegeven in één grafiek (figuur 3). Vervolgens zijn per uur de 10^e, 50^e en 90^e percentielen bepaald (figuur 4). Tussen de 10^e en 90^e percentielen zijn 80% van de waargenomen uurintensiteiten gelegen. Daarnaast is voor ieder cluster de variatiecoëfficiënt² bepaald tijdens de ochtendspits en de avondspits. Hiervoor zijn de absolute spitsperiodes gebruikt. Tot slot is het tijdstip van de absolute en relatieve spitsperiode vergeleken voor de dagen binnen een cluster.



Figuur 3: variatie tussen dagen binnen clusters



Figuur 4: 10^e, 50^e en 90^e percentielen

De dagpatronen binnen de verschillende clusters blijken behoorlijk te variëren. Met name clusters 2 en 3 bevatten dagen waarvan het dagpatroon verschillende vormen vertoont. Binnen cluster 1 verschillen de dagpatronen minder qua vorm. Ook de verschillen in uurintensiteiten tussen verschillende dagen zijn kleiner, resulterend in een smallere bandbreedte in figuur 2. In cluster 1 is echter wel een dag met een duidelijk hogere intensiteit gedurende bijna de hele dag. Waarschijnlijk is dit het gevolg van een evenement.

	Ochtendspits	Restdag	Avondspits
Cluster 1	3.58	6.45	6.12
Cluster 2	4.96	7.19	5.69
Cluster 3	5.65	9.42	7.77

Tabel 2: variatiecoëfficiënt op verschillende tijdstippen

² $(100 * \text{standaardafwijking}) / \text{gemiddelde}$

De variatiecoëfficiënt is tijdens de ochtendspits lager dan tijdens de rest van de dag. Daarnaast blijkt ook uit tabel 2 dat de variatie binnen cluster 1 kleiner is dan de variatie binnen clusters 2 en 3.

Wanneer de spittijden met elkaar vergeleken worden, blijkt dat eenderde van de dagen in cluster 3 een zeer late ochtendspits vertoont. De twee opeenvolgende drukste uren liggen op deze dagen tussen 10:00 en 12:00 uur in plaats van de normale spittijd tussen 7:00 en 9:00. De twee drukste uren in de middag variëren zowel voor dagen binnen cluster 2 (15:00 – 17:00 en 16:00 – 18:00) als cluster 3 (14:00 – 16:00 en 16:00 – 18:00). De relatieve spitsperioden zijn voor geen enkel cluster voor alle dagen precies gelijk. De ochtendspits start voor alle dagen in cluster 1 om 7:00 uur, maar duurt in sommige gevallen tot 8:00 uur in plaats van 9:00 uur. De avondspits start tussen 14:00 uur en 16:00 en duurt tot 18:00 of 19:00. Voor dagen binnen cluster 2 is de variatie iets groter. De ochtendspits start bijna altijd om 7:00, maar op één dag om 8:00 uur en duurt soms tot 8:00 uur en soms tot 9:00 uur. De avondspits begint tussen 13:00 uur en 16:00 en duurt tot 17:00 uur, 18:00 uur of 19:00 uur. Cluster 3 kent duidelijk de grootste variatie. Bijna de helft van de dagen vertoont geen echte ochtendspits. Op de andere dagen begint de ochtendspits om 7:00 of om 10:00 uur en eindigt de ochtendspits om 8:00 uur, 9:00 uur, 11:00 of 12:00 uur. De avondspits start tussen 13:00 uur en 16:00 uur en is afgelopen tussen 17:00 en 19:00.

Om de verschillen tussen dagen binnen clusters te verkleinen kan gekozen worden voor een indeling in meerdere clusters. Volgens het dendogram is een indeling in 5 clusters ook mogelijk. In dit geval wordt cluster 2 gesplitst in twee kleinere clusters met ieder 8 dagen en wordt de afwijkende dag uit cluster 1 ook als apart cluster meegenomen. Ook bij deze clustering blijven echter aanzienlijke verschillen tussen dagen binnen clusters bestaan. Een getrapte clustering op basis van afzonderlijke kenmerken van het dagpatroon leidt wellicht tot een betere clustering.

Nadere analyse ontstane clusters

Allereerst is voor de dagen binnen de verschillende clusters bepaald of ze al dan niet in de bouwvak- en de schoolvakantie gelegen zijn. Alle dagen in cluster 3 blijken binnen de

schoolvakantie te liggen en slechts 1 van de 13 dagen ligt buiten de bouwvakvakantie. Alle dagen in clusters 1 en 2 liggen daarentegen buiten de bouwvakvakantie. Bijna tweederde van de dagen in cluster 2 ligt echter wel in de schoolvakantie. Van de dagen in cluster 1 ligt 85% buiten de schoolvakantie. Geconcludeerd kan dan ook worden dat met name de bouwvakvakantie een grote invloed heeft op de clustering.

	Maandag	Dinsdag	Donderdag	Vrijdag
Cluster 1	9	9	9	0
Cluster 2	2	3	2	9
Cluster 3	4	3	3	3

Tabel 3: Verdeling van dagen van de week over de verschillende clusters

Het meest opvallende verschil met betrekking tot het aantal verschillende wekdagen (zie tabel 3) is waar te nemen tussen clusters 1 en 2. Cluster 1 bevat geen enkele vrijdag, terwijl meer dan de helft van de dagen binnen cluster 2 een vrijdag is. Wanneer alleen dagen buiten de vakantie geselecteerd worden blijkt zelfs dat alle vrijdagen buiten de vakantie in cluster 2 gelegen zijn en cluster 2 alleen vrijdagen bevat. Vrijdag blijken dus in een ander cluster te worden ingedeeld dan andere werkdagen.

5. Conclusie

In steeds meer steden komt –mede als gevolg van de implementatie van stedelijke verkeerscentrales- verkeersinformatie beschikbaar. Met behulp van informatie die over een langere periode is ingewonnen kunnen stedelijke verkeerspatronen gedefinieerd worden die vervolgens gebruikt kunnen worden voor stedelijk verkeersmanagement en verkeersvoorspellingen. Dit paper beschrijft een voorstel voor de analyse van stedelijke dagen en spitspatronen.

Om de beschreven procedure te illustreren is een clustering uitgevoerd op een kleine hoeveelheid data. Dagpatronen zijn ingedeeld in clusters met behulp van Ward's method. Dagpatronen zijn daarbij gedefinieerd als een serie intensiteiten (in dit geval uurintensiteiten) als functie van de tijd. De clustering blijkt tot een indeling in duidelijk van elkaar te onderscheiden clusters te leiden. De gemiddelde dagpatronen van de afzonderlijke clusters verschillen zowel in hoogte als in vorm. De variatie tussen dagen binnen de clusters blijkt

echter wel behoorlijk te zijn. In twee van de drie ontstane clusters variëren de dagpatronen onderling aanzienlijk qua hoogte en qua vorm. Dit is niet gewenst met betrekking tot verkeersmanagement- en voorspellingsdoeleinden. Een indeling in meer clusters blijkt niet tot een beter resultaat te leiden. Het verdient dan ook aanbeveling om de resultaten bij een andere definitie van een dagpatroon te analyseren. Wanneer een dagpatroon gedefinieerd wordt aan de hand van een aantal kenmerken die samen de hoogte en vorm van het patroon vastleggen, ontstaat wellicht een clustering waarbij de dagen binnen clusters beter overeenkomen met betrekking tot deze kenmerken. Kenmerken die van belang zijn met betrekking tot verkeersmanagement zijn de spitsintensiteiten, de totale intensiteit, de spitsstijden en de verhoudingen tussen dagdelen. Door middel van een getrapte clustering kunnen de dagen op basis van deze kenmerken in groepen worden ingedeeld. Omdat de dataset slechts een beperkt aantal dagen bevatte, kon deze methode nog niet worden uitgetoetst.

Tot slot is onderzocht welke factoren ten grondslag liggen aan de ontstane clusters. In deze dataset konden twee factoren worden meegenomen, namelijk de dag van de week³ en het al dan niet onderdeel uitmaken van een vakantieperiode. Vakantieperioden bleken het dagpatroon behoorlijk te beïnvloeden. Alle dagen binnen de bouwvakvakantie zijn in eenzelfde cluster ingedeeld. Dit cluster vertoont de gehele dag een lagere intensiteit dan de andere clusters. Ook vrijdagen buiten de vakantie zijn allemaal in hetzelfde cluster ingedeeld. Dit cluster vertoont een vlakker dagpatroon dan het cluster met maandagen, dinsdagen en donderdagen. Wel moet hierbij opgemerkt worden dat beide clusters ook (vrij)dagen binnen de schoolvakantie bevatten.

³ De dataset bevatte gegevens van maandagen, dinsdagen, donderdagen en vrijdagen

Referenties

- Ancidei, V. Cera, E., De Pisi, P. Francalanci, S. Landolfi, O. & Tomassini, M. (2000). Integrated ITS in the city of Rome. *Proceedings of the 7th World Congress on Intelligent Transport Systems*, Turin 2000.
- Bae, S. & Lee, B. G. (2000). A real-time traffic information service by dedicated fm broadcasting system in Seoul Korea. *Proceedings of the 7th World Congress on Intelligent Transport Systems*, Turin 2000.
- Barcélo, J. (2000). The role of simulation in advanced traffic management systems. *Proceedings of the 7th World Congress on Intelligent Transport Systems*, Turin 2000.
- Budde, A. (2002). Düsseldorf's traffic management system. *Proceedings of the 9th World Congress on Intelligent Transport Systems*, Chicago, 2002.
- Clarke, R. & Morris, R. (2001). Integration of control and information systems for transport providing a strategic co-ordinated approach. *Proceedings 8th World Congress on Intelligent Transport Systems*, Sydney 2001.
- Cone, R. James, L. Jeffery, D. Fisher, G. & Wainwright, A. (2002) Delivering ITS and travel information in Wales. *Proceedings of the 9th World Congress on Intelligent Transport Systems*, Chicago, 2002.
- Fellendorf, M., Nökel, K. & Handke, N. (2000). VISUM-online –traffic management for the EXPO 2000 based on a traffic model. *Proceedings of the 7th World Congress on Intelligent Transport Systems*, Turin 2000.
- Ferulano, G., Gortan, L., Sforza, A. & Tartaro, D. (2000). The ATENA project “Ambiente traffico telematica Napoli”. *Proceedings of the 7th World Congress on Intelligent Transport Systems*, Turin 2000.
- FHWA (2003). Traveler Information Systems in Europe. FHWA-PL-03-005.
http://international.fhwa.dot.gov/travelinfo/traveler_information.pdf
- Hasberg, P. & Serwill, D. (2000). Stadtfoköln –a global mobility information system for the Cologne area. *Proceedings of the 7th World Congress on Intelligent Transport Systems* Turin.
- Henriet, E. & Schmitz, P. (2000). Information strategies in the urban context. *Proceedings of the 7th World Congress on Intelligent Transport Systems*, Turin 2000.
- Hoyer, R. & Herrmann, A. (2003). ITS in small innovative steps –successful deployment in a medium-sized city. *Proceedings of the 10th World Congress on Intelligent Transport Systems*, Madrid, 2003.

Jain, A.K., Murty, M.N. & Flynn, P.J. (1999). Data Clustering: A Review. *ACM Computing Surveys*, **31(3)** pp 264 – 323.

Karl, C.A. & Trayford, R.S. (2000) Delivery of real-time and predictive travel time information: Experiences from a Melbourne trial. *Proceedings of the 7th World Congress on Intelligent Transport Systems*, Turin 2000.

Kellerman, A. & Schmid, A. (2000) Mobinet: Intermodal traffic management in Munich – control centre development. *Proceedings of the 7th World Congress on Intelligent Transport Systems*, Turin 2000.

Kirschfink, H. Riegelhoth, G. & Boero, M. (2000). Strategic management by the enterprise approach. *Proceedings 7th World Congress on Intelligent Transport Systems*, Turin 2000.

Kruse, G., Tannert, R. & Hasberg, P. (2000). Incident detection by MOTION for strategic control in the traffic management system “stadtfoköln”. *Proceedings of the 7th World Congress on Intelligent Transport Systems*, Turin 2000.

Leitsch, B. (2002) A Public-private partnership for mobility – Traffic Management Center Berlin. *Proceedings of the 9th World Congress on Intelligent Transport Systems*, Chicago.

MacLennan, C. & Routledge, I.W. (1995). The development of Urban Traffic Management and Control (UTMC) systems. *Traffic Engineering and Control* **36 (9)** pp. 468-471.

Manchester Metropolitan University, *Cluster Analysis* <http://149.170.199.144/multivar/ca.htm>

MIZAR Automazione (1998); COSMOS (Congestion Management Strategies and Methods in Urban Sites) *report on Torino’s Demonstrator for Congestion and Incident Detection and Management*, deliverable D04.3, Telematics Application Program Transport (TR1015).

Nowotny, B., Asamer, J., Kashif, D. & Karim, R. (2003). Classification of traffic data time series by cluster analysis, artificial neural networks and ANOVA. *Proceedings of the 10th World Congress on Intelligent Transport Systems*, Madrid, 2003.

Richards, A. Piao, J. & McDonald, M. (2000). VMS in Southampton: A case study. *Proceedings of the 7th World Congress on Intelligent Transport Systems*, Turin 2000.

Scharrer, R., Kippes, G., Glas, F. & Keller, H. (2003). An Innovative Road Side Driver Information System –NetzInfo-, *Proceedings of the 10th World Congress on Intelligent Transport Systems*, Madrid, 2003.

Webb, A.R. (2002). *Statistical pattern recognition*. Chichester: Wiley.

Weijermars, W.A.M. & E.C. van Berkum (2004). Daily flow profiles of urban traffic. In: eds. Brebbia, C.A. & Wadhwa, L.C. *Urban Transport X*, pp. 173-182. Southampton: WITpress.

Wild, D. (1997). Short-term forecasting based on a transformation and classification of traffic volume time series. *International Journal of Forecasting* **13**, pp 63-72.