datamanagementgroup

2015 DMG Annual Report

Prepared by:

Data Management Group
Department of Civil Engineering
University of Toronto

INTRODUCTION

The Data Management Group (DMG) was established in 1988 on the basis of a proposal from the University of Toronto's Joint Program in Transportation for an autonomous research group with the following objectives:

(a) to establish a common, centrally-accessible database containing information on transportation activities, zone systems, transportation networks and land use activity,
(b) to provide a transportation data retrieval service to the participating agencies,
(c) to monitor the adequacy of available data and propose approaches for adding to or updating the data as mutually agreed upon by the agencies,
(d) to promote greater interaction between university researchers and practitioners in the field of urban transportation planning,
(e) to promote the communication of transportation information and data obtained or administered by the Data Management Group to interested agencies and to the public, and
(f) to further the improvement of transportation demand analysis, research, and forecasting in the Greater Toronto Area.

Although the administration of the group has changed to the Department of Civil Engineering at the University of Toronto, the DMG continues to be guided by these objectives into this its 26th year of continuous operation.

Program approval and funding of the DMG is the collective responsibility of members of the Transportation Information Steering Committee (TISC) with the following membership:

City of Hamilton
City of Toronto
Metrolinx
Ministry of Transportation, Ontario
Regional Municipality of Durham
Regional Municipality of Halton
Regional Municipality of Peel
Regional Municipality of York
Toronto Transit Commission

Each participating agency appoints a member of their technical staff to the Transportation Research and Data Management Group (TRADMAG), which is a

standing committee of TISC, and is responsible for coordinating the needs of the funding agencies and the activities of the research project.

This report provides a brief profile of the staff employed and a description of the activities undertaken by the DMG during the calendar year 2015.

STAFF AND LOCATION

Director

Eric J. Miller, Professor, Department of Civil Engineering, University of Toronto

Technical Staff

Reuben Briggs, B.A.Sc. (Civil Engineering), M.A.Sc. (Civil Engineering) University of Toronto, P.Eng.

Jason Chan, B.A.Sc. (Electrical & Computer Engineering) University of Toronto

Susanna Choy, B.A.Sc. (Industrial Engineering), M.Eng. (Civil Engineering) University of Toronto, P.Eng.

Urban Transportation Planning Intern

Desmond Chan, B.A.Sc. (Civil Engineering) University of Toronto (May 2014 - April 2015)

Web Site

http://www.dmg.utoronto.on.ca

Office Location

Department of Civil Engineering
University of Toronto
Galbraith Building, Suite 305
35 St. George Street
Toronto, Ontario M5S 1A4
Telephone: (416)978-3913
Email: info@dmg.utoronto.ca

TRANSPORTATION TOMORROW SURVEY

Under the guidance of TISC, a series of urban travel surveys, Transportation Tomorrow Survey (TTS), have been conducted every five years since 1986. The last TTS was conducted in 2011/2012 and was managed by the Data Management Group in a manner similar to all previous TTS since 1991.

Due to the absence of a principal investigator from the University of Toronto, the DMG cannot undertake the overall administrative responsibility for the 2016 TTS. As a result, TISC will issue a contract to an external consultant through the Ministry of Transportation's procurement office for the delivery of the 2016 TTS. The DMG presented to TISC the cost estimates and requirements for the 2016 TTS and prepared the request for proposal document. This document describes the scope of the survey, the coverage area, detailed procedures involved in the survey, quality control measures, expected deliverables and timelines for major tasks. It was announced on the Ontario Government e-Tendering system in November 2015. The DMG also created several survey software documents and demonstration videos for the potential vendors as it is a requirement to conduct the survey with the same software used in 2011 TTS. An information session was held by the Ministry of Transportation and DMG in December to provide information on the procurement process and survey procedures to potential vendors. A similar information session on software demonstrations would be held in mid January 2016. This would be followed by software review sessions provided to each potential vendor by the DMG. Each vendor could send two staff members to DMG to test the Direct Data Entry and Geocoding modules of the software on a designated computer at the DMG office.

The deadline for the submission of proposals was January 29th, 2016. The DMG would be part of the evaluation committee to review and assess the submitted proposals. The award of the contract is expected to occur in April 2016.

INFORMATION PROCESSING

The term 'information processing' is used in this instance to describe a set of activities supporting the management, storage and distribution of urban travel information. The principle components of this information are the results of the Transportation Tomorrow Surveys and a collection of all Cordon Count Surveys.

Transportation Tomorrow Surveys and iDRS

The DMG administers the data files on urban travel contained in the 1986, 1991, 1996, 2001, 2006 and 2011 Transportation Tomorrow Surveys (TTS) in the form of a set of relational databases with various methods of access. Direct access to the original files is restricted to DMG staff to ensure that information on a particular household cannot be identified. Currently, data files available to iDRS users contain the following information:

TTS Number of Records

Year	Households	Persons	All Trips	Transit Detail
1986	61,453	171,086	370,248	56,615
1991	24,507	72,496	157,349	14,896
1996	115,193	312,781	657,971	70,295
2001	136,379	374,182	817,744	85,095
2006	149,631	401,653	864,348	87,244
2011	159,157	410,404	858,848	86,703

The increasing size of the databases reflects not only growth in the area but also changes in the size of the area surveyed.

The first text-based data retrieval system (drs) was developed in the mid 1990s by the staff of DMG as the original method for external users to gain access to the data files and complete the data extraction themselves. This retrieval system was very effective when a modem was used as the principle method of remote access to the DMG's computer system. Over the years, as the demand for travel data grew and the Internet became the preferred method of remote access, a data retrieval system specifically designed for Internet access was developed (iDRS). All the features of drs were incorporated into the browser-based iDRS and the drs process was phased out.

The initial release of iDRS was restricted to use by the funding agencies. In 2002, access to iDRS was made available to any individual that requested access. The individual is required to sign an agreement form and system security is maintained by giving each user a unique login and password. This procedure has the added benefit

that agencies outside the Greater Toronto and Hamilton Area that participated in the Transportation Tomorrow Surveys can access their data without the need to set up their own database system.

New iDRS

In order to meet the changing needs of the data, the DMG has developed a new version of iDRS. The software was rewritten to incorporate a new database management system PostgreSQL, which was used as part of the TTS survey software. Several new functions such as additional filter criteria and ability to store the query for future use have been added and the speed of extraction has increased dramatically.

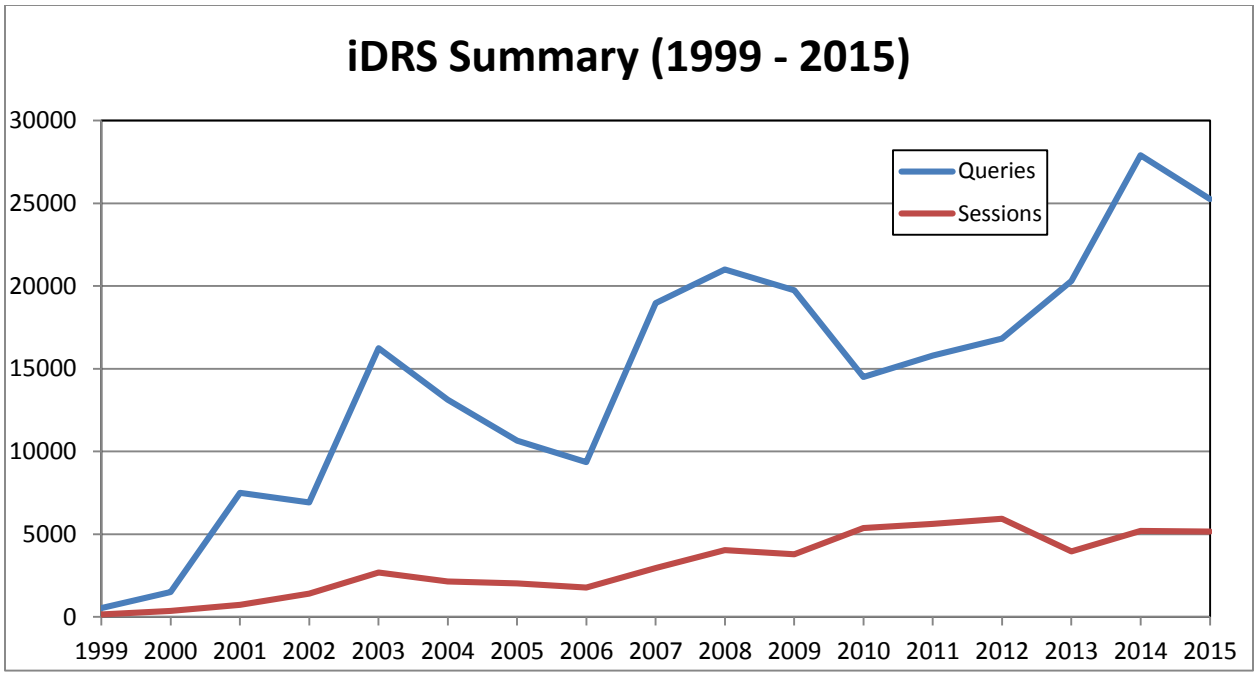
A new authorization process has also been implemented together with the new software. The process has been automated so that any individual can request access by entering the name, affiliation, email address and contact number online. Once the information is submitted, it will be approved by the DMG staff and an email with the credentials will be sent to the new user.

The beta version of the new iDRS was released in September 2015. There were 42 external users selected to test the software and 519 queries were generated. The beta version only allows TTS data extraction. The DMG continues to fine-tune the new iDRS based on the feedback of the selected users and anticipates to release it officially in 2016.

A History of iDRS Data Requests

When the browser based data extraction procedure (iDRS) was first released in 1999, the users were registered users of the DMG's computer system. The initial use of iDRS was encouraging. Subsequently in late 2000, the Steering Committee gave approval for use by anyone registering and providing information on their affiliation. The growth in use of TTS data is reflected in the growth of the use of iDRS for data extraction. The following chart shows the growth since iDRS was first introduced in 1999. The 'Number of Sessions' reflects the number of times registered users, including DMG staff, have initiated a data retrieval session. The 'Number of Queries' reflects the number of times an output was generated during a session. Almost without exception, many queries are generated during a given session. Note that the

iDRS usage peaked at the year after a new TTS dataset was released in 2002, 2007 and 2014[1].



Summary of iDRS Data Requests in 2015

In 2015, there were 25,263 iDRS queries within 5,156 sessions. The following table shows the monthly summaries.

Month	Queries	Sessions
January	3262	628
February	3106	694
March	2393	616
April	1871	358
May	1460	366
June	2006	387
July	4588	655
August	1586	366
September	1174	263
October	1422	354
November	1243	272

[1] The interview phase of the 2011 TTS completed in December 2012 and the final data set was released in November 2013.

December	1152	197

Total 2015	**25263**	**5156**
Total 2014[2]	27901	5194

Affiliations of iDRS Users in 2015

The following is a list of 80 different agencies and groups that extracted the data through iDRS in 2015.

407 ETR	AECOM Canada
Artsape	Arup Canada
BA Consulting Group Ltd.	Caliper Corporation
Cancer Care Ontario	CF Crozier & Associates
City of Brampton	City of Burlington
City of Guelph	City of Hamilton
City of Markham	City of Mississauga
City of Peterborough	City of Toronto
City of Vaughan	Cole Engineering Group Ltd.
Community Development Halton	Community Reach
Concordia University	Consumer Policy Institute
CPCS Transcom Ltd.	David Kriger Consultants Inc.
Delcan Corp.	Dillon Consulting Ltd.
Dionne Bacchus & Associates	Entra Consultants
Environics Analytics	EXP services
GHD Inc.	GreenUP (Peterborough)
H Fox Associates Ltd.	Hatch Mott MacDonald
HBA Specto Inc.	HDR Inc.
IBI Consulting Group	JD Northcote Engineering Inc.
LEA Consulting Inc.	Lee-Gosselin Associates
McGill University	McMaster University
Meliane Inc.	Metrolinx
Ministry of Transportation Ontario	Mohawk College
Neptis Foundation	Paradigm Transportation Solutions
Pembina Institute	Poulos & Chung Ltd.
Queen's University	R. J. Burnside & Associates Ltd.
Read Voorhees and Associates Ltd.	Region Municipality of Durham

[2] The number of iDRS queries and sessions in 2014 have been adjusted to exclude extractions to generate results for Open Data.

Region Municipality of Halton	Region Municipality of Niagara
Region Municipality of Peel	Region Municipality of Waterloo
Region Municipality of York	RideCo
Ryerson University	Sernas Group Inc.
Social Planning and Research Council of Hamilton	Stantec Architecture Ltd.
Steer Davies Gleave	Sustainable Urban Development Association
Tedesco Engineering	The Arland Group Ltd.
The Atmospheric Fund	The Record
Toronto Cycling Think and Do Tank	Toronto Transit Commission
Town of Ajax	Town of Newmarket
Town of Richmond Hill	Tri-Cities Transport Action Group
UEM Consulting	University of Toronto
WSP Canada Inc.	York University

TTS Special Data Requests

The interactive procedures available with iDRS satisfy the majority of data needs. However, some data needs are too complex and require the intervention of an experienced analyst to formulate a custom query from the database. In addition, the DMG staff can often help define the most relevant data for the problem at hand. One of the typical data requests involves assigning the x-y coordinates of households or trip ends to a specified spatial aggregation different from the predefined traffic zone system as these coordinates are not available to any non DMG staff in order to protect the privacy of the survey respondents. There is also a growing interest in using disaggregate data for travel forecasting among academic researchers. Here is a list of agencies that have made a special data request in 2015:

City of Guelph
City of Toronto
McMaster University
Metrolinx
Ministry of Transportation, Ontario
Regional Municipality of Niagara
University of Cambridge
University of Toronto

Cordon Counts and CCDRS

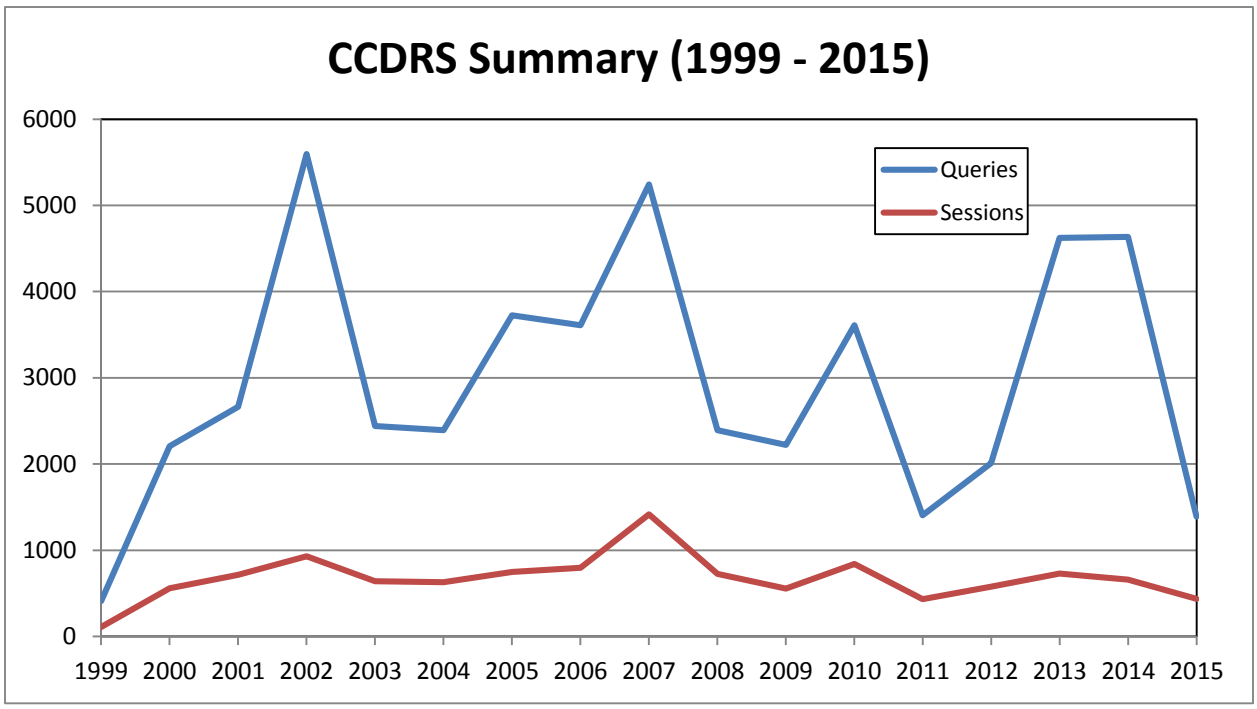
The City of Toronto (then the Regional Municipality of Metropolitan Toronto) began collecting detailed information on the type and volume of traffic crossing selected points on the road system as early as 1975. The counting locations were selected such that screen lines or cordon lines could be defined and the counting program has continued on a regular basis since that time, usually twice in a five year cycle. Subsequently, other Regions began similar programs. Given the number of Regions with a similar program, they began coordinating their count programs and defining a common set of data standards. In 1998, the DMG collected these traffic counts in a common database structure and developed a Cordon Count Data Retrieval System (CCDRS). CCDRS is now widely used by a variety of public and private agencies.

In addition, participating agencies now use CCDRS as a tool in verifying their cordon count results. The approved procedure is to place new cordon count data directly into the CCDRS database with a disclaimer notice to all users that the new information is preliminary. The last Cordon Count was undertaken by the participating regions in 2014.

The DMG has started the development of the new CCDRS. The final product will include a graphic interface so the users can identify and select screen lines and count stations from a map. The beta version of the new CCDRS is anticipated to be release in 2016.

A History of CCDRS Data Requests

The growth in access to the CCDRS data is reflected in the increased number of data extractions. The following chart shows the growth since CCDRS was first introduced in 1999. The 'Number of Sessions' reflects the number of times registered users, including DMG staff, have initiated a data retrieval session. The 'Number of Queries' reflects the number of times an output was generated during a session. Almost without exception, many queries are generated during a given session. The busiest years were just after restrictions on access to the data were removed. An increase in activity is usually associated with the release of a new cordon count. The 2014 cordon counts were partially released at the end of 2015 with data from the City of Toronto, the Regions of Peel and Halton. The remaining data from the Regions of Durham and York were released in early 2016.



Summary of CCDRS Data Requests in 2015

In 2015, there were 1,392 CCDRS queries within 436 sessions. The following table shows the monthly summaries.

Month	Queries	Sessions
January	193	48
February	150	45
March	161	58
April	84	19
May	141	47
June	124	40
July	43	26
August	97	36
September	118	37
October	42	22
November	142	31
December	97	27
Total 2015	**1392**	**456**
Total 2014[3]	4633	658

Affiliations of CCDRS Users in 2015

The following is a list of 31 different agencies and groups that extracted the data through CCDRS in 2015.

AECOM	BA Group
City of Brampton	City of Burlington
City of Toronto	Cole Engineering Group Ltd.
David Kriger Consultants Inc.	Dillon Consulting Ltd.
EXP Services	H Fox Associates Ltd.
HDR Inc.	IBI Consulting Group
LEA Consulting Inc.	Leighfisher Ltd.
Meliane Inc.	Metrolinx
Ministry of Transportation Ontario	Morrison Hershfield Ltd.
Paradigm Transportation Solutions Ltd.	Parsons Corp.
Pembina Institute	Peter Dalton Consulting
Region Municipality of Durham	Region Municipality of Peel
Region Municipality of York	Ryerson University
SNC-Lavalin Group Inc.	University of Guelph
University of Toronto	Valcoustics Canada Ltd.
WSP Canada Inc.	

[3] The number of 2014 CCDRS queries and sessions have been adjusted to exclude extractions to generate the results for Open Data.

CCDRS Special Data Requests

The interactive procedures available with CCDRS satisfy the majority of data needs. However, some data needs are too complex and require the intervention of an experienced analyst to formulate a custom query from the database. In addition, the DMG staff can often help define the most relevant data for the problem at hand. Although special data requests are an important function, an objective of the DMG is to reduce the number of such data requests in favour of users processing their request through CCDRS. There were no special data request related to the cordon counts in 2015.

Open Data

In March 2011, the first generation of Open Data Portal was launched by the Government of Canada. "Open Data is defined as structured data that is machine-readable, freely shared, used and built on without restrictions." The data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the internet. It must also be available in a convenient and modifiable form and must be provided under terms that permit re-use and redistribution including the intermixing with other datasets. Everyone must be able to use, re-use and redistribute. There should be no discrimination against fields of endeavour or against persons or groups. For example, 'non-commercial' restrictions that would prevent 'commercial' use, or restrictions of use for certain purposes (e.g. only in education), are not allowed.

With the approval from TISC, the DMG has opened up the TTS and cordon count data to the public in 2014. Unlike the online data retrieval systems, users can download the pre-generated text files from the DMG web site without registration. There are currently 103 TTS data files and 204 Cordon Count data files stored in the Open Data Portal at the DMG. Each TTS data file contains household, person, and trip information for a specific survey year, geographic area and spatial aggregation. Each cordon count data file contains different types of vehicle and person counts for a specific year, geographic area and time period. The files are in comma-delimited text format and readily import into Excel or other spreadsheet software. The DMG continues to update the Open Data Portal to include new TTS and cordon count data files.

COMPUTER RESOURCES AND TECHNICAL SUPPORT

The DMG computer system has been redesigned to improve the performance and system migration. It is currently comprised of several servers including the DMG main server, the data retrieval server, an EMME2 server, an EMME3/EMME4 licence server, a backup server and a file server where the data files of the funding agencies are stored. All these servers are located behind the DMG firewall for security reasons.

The DMG main server is a Dell R620 running Windows Server 2012 R2 with Hyper-V application. This application allows the DMG server to run multiple virtual machines, i.e. operating systems emulated within another operating systems by imitating dedicated hardware. There are three virtual machines running on the DMG server. The first one is the mail server using Microsoft Exchange Server 2013 to handle daily DMG emails. The second one is the DMG website created in WordPress. Linux, Apache, MySQL, PHP and Lamp stack are running on this machine. The third virtual machine is the beta version of the new iDRS developed using the MEAN stack, a free and open source JavaScript software stack for building web application running in Debian. The data are hosted in PostgreSQL which is an open source relationship database management system. The use of open source software allows free access, modification and sharing of the software. The beta iDRS is running simultaneously with the legacy system. Since June 2015, new users have only been allowed to access the new beta iDRS in an effort to test its capabilities and also to facilitate the full switch-over from the old iDRS which is expected in 2016.

The data retrieval server hosts the TTS, Cordon Count, 1964 MTARTS and GO Transit survey (prior to 1997) databases. These data are stored in Oracle, a relational database system, on a SUN Ultra Sparc server running Solaris and can be accessed through the two interactive data retrieval systems iDRS or CCDRS. All years of the TTS and Cordon Count Data are compiled and stored in a consistent way, and the TTS and Cordon Count data are frequently used together in regional planning projects or traffic impact studies. A summary of the use and the users of these services is contained earlier in this annual report.

The EMME2 server allows funding agencies and their consultants to run an older version of EMME2 remotely on a Sun server running Solaris. EMME2 is no longer supported by INRO, the developer of the EMME software and most of the agencies have already migrated their modelling efforts to EMME3 or EMME4.

The newer EMME3/EMME4 software is run locally on a personal computer with all relate files stored locally. This results in much faster operation. Each local machine requires access to an authentication key to operate the software. Rather than each agency purchasing a licence from INRO to use the software locally, DMG negotiated a

concurrent licence for authentication of several machines operating at the same time. The DMG then dedicated a server to provide remote authentication to the participating partners. The DMG is in the process of migrating the current suite of supporting hardware and software to more recent and cost effective alternatives.

In 2015 DMG also upgraded the network infrastructure from 200 megabits per second to 10 gigabits per second, by updating the firewall and installing a fiber cable directly to the router. This infrastructure upgrade will significantly speed up the throughput of all the entire computer systems: iDRS, CCDRS, EMME proxy, and the DMG website.

For the 2016 survey, DMG purchased a Dell SonicWall TZ250 Firewall and two Dell PowerEdge R430 servers. The R430s servers will be running on virtual machines with redundancies for the survey, like our main system. Configuration of these new servers was started in 2015.

The DMG continues to maintain and improve the computer system to meet its funding agencies' changing needs.