Evaluating Long-Distance Travel Patterns in Israel by Tracking Cellular Phone Positions

Journal:	*Journal of Advanced Transportation*
Manuscript ID:	Draft
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Bekhor, Shlomo; Technion - Israel Institute of Technology, Civil and Environmental Engineering Cohen, Yehoshua; Matat – Transportation Planning Center Ltd Solomon, Charles; Ministry of Transport and Road Safety, Economics and Planning Department
Keywords:	Intelligent transport systems, Transport planning, Transportation analysis
Abstract:	Long-distance trips are generally under-reported in typical household surveys, because of relative low frequency of these trips. This paper proposes to utilize location data from cellular phone systems in order to study long-distance travel patterns. The proposed approach allows passive data collection on many travelers over a long period of time at low costs. The paper presents the results of a study that applies cellular phone technology to assess trips at the national level. The method was specifically designed to capture long distance trips, as part of the development of a national demand model conducted for the Economics and Planning Department of the Israel Ministry of Transport. The method allows the construction of Origin-Destination tables directly from the cellular phone positions. The paper presents selected results to illustrate the potential of the method for transportation planning and analysis.

Evaluating Long-Distance Travel Patterns in Israel by Tracking Cellular Phone Positions

Shlomo Bekhor, Yehoshua Cohen, Charles Solomon

1. Associate Professor, Transportation Research Institute, Technion City, Haifa 32000, Israel

Tel: 972-4-8292460, Fax: 972-4-8225716, Email: sbekhor@technion.ac.il

2. Matat – Transportation Planning Center Ltd., 43 Haaliyah Hashniyah St., POB 189, Azor 58190, Israel

Tel: 972-3-5588850, Fax: 972-3-5587965, Email: matat@netvision.net.il

3. Deputy Director General, Economics and Planning Department, Ministry of Transport and Road Safety, 5 Bank Israel St., Jerusalem 91950, Israel

Tel: 972-2-6663080, Fax: 972-2-6663083, Email: charles@mot.gov.il

Evaluating Travel Demand Patterns in Israel by Tracking Cellular Phone Positions

ABSTRACT

Long-distance trips are generally under-reported in typical household surveys, because of relative low frequency of these trips. This paper proposes to utilize location data from cellular phone systems in order to study long-distance travel patterns. The proposed approach allows passive data collection on many travelers over a long period of time at low costs. The paper presents the results of a study that applies cellular phone technology to assess trips at the national level.

The method was specifically designed to capture long distance trips, as part of the development of a national demand model conducted for the Economics and Planning Department of the Israel Ministry of Transport. The method allows the construction of Origin-Destination tables directly from the cellular phone positions. The paper presents selected results to illustrate the potential of the method for transportation planning and analysis.

1. INTRODUCTION

Transportation analysis and travel demand forecasting relies on survey data to provide inputs and to calibrate the mathematical equations that represent decisions people make related to travel. Literature shows that the primary data collection method for travel demand modeling is household surveys, often complemented by other surveys such as traffic and person counts, public transport on-board surveys, intercept surveys at external cordon points, commercial vehicle surveys, and others. According to Stopher et al. (2006), household travel surveys continue to be a mainstay of transport planning and modeling efforts, despite of well-known problems that are briefly described.

A problem discussed in the literature, common to all interview-type surveys, is non-response. Complex methods to correct for non-response have been developed, however these alleviate the problem only partially (Richardson and Wolf, 2001). Another limitation of household surveys is the need for active cooperation from the respondents, relying on their memory and patience. The need for active participation reduces the ability to capture complex travel and activity patterns, and the ability to collect data over a long period of time. The problems mentioned above, coupled with budget constraints, explain the fact that typical household surveys collect data about 1 or 2 days for each household.

The potential of long-term surveys for understanding regularities and deviations in travel behavior has been explored by Axhausen et al. (2004), who analyzed a six-week continuous travel diary performed by the German Ministry of Research and Education. The authors stress the importance of collecting data over long duration to capture effects of transport policy measures on the dynamic processes in travel behavior. Although the findings of this important effort are very insightful, their method for data collection is rather costly, and not likely to be repeated in other countries, at least not very often.

To alleviate the problems of memory and precision, recent travel survey are supplemented by GPS traces of personal and vehicular travel (Richardson and Wolf, 2001). Typically, the GPS device is provided to respondents by the survey organizers. The scope of GPS supplemented surveys is often rather limited, due to operational considerations and particularly the cost of the GPS devices. Even with reduction of GPS costs, there are still high costs related to the fact that the GPS must be handed to each household, together with some explanation on how to use it, and subsequently retrieving the data.

Recent developments offer in addition to GPS another alternative Wireless Location Technology (WLT), which is based on cellular phones. Cellular networks are based on the division of the service area into small geographic units (cells), served by different antennas. The fundamental system design requires the ability to locate every cellular phone, in order to be able to provide the service. WLT is used in a wide range of commercial applications, aimed at providing cellular phone customers with location-based services, such as information about nearby facilities of interest (gasoline stations, restaurants, etc.); route guidance to a chosen destination; traffic conditions alerts; and more.

In Israel, there is a lack of information about long-distance trips. Although national household surveys were performed, the data collected about long-distance trips could neither be used to estimate aggregate Origin-Destination matrices, nor to estimate trip distribution models. Given the high penetration of cellular phones in Israel, an alternative solution was proposed. The idea was to utilize location data from cellular phone systems in order to study long-distance trips.

This paper presents the results of a study that applies cellular phone technology to assess trips at the national level. To our knowledge, this is one of the first attempts performed at this scale for transportation planning, and it encompasses large amounts of spatial and temporal data. The method was designed specifically to capture long distance trips, as part of the development of a national demand model, which is being conducted for the Economics and Planning Department of the Israel Ministry of Transport.

The proposed approach allows collection of data on many travelers at high response rates over a long period of time at low costs. The approach is expected to achieve high temporal accuracy and reasonable spatial accuracy. Furthermore, cellular phones based data avoid the problems related to the reliance on respondents' memory. While the approach has its limitations, as discussed later in this paper, it has the potential of generating a significant change in the abilities to analyze, understand, and predict travel behavior.

A new innovative data collection methodology introduces many risks and uncertainties. The expected significance of this paper is to provide an empirically based assessment of the potential of this data collection method, with quantified evaluation of its advantages and limitations.

2. LITERATURE REVIEW

Qiu and Cheng (2007) review twenty studies worldwide, in which cellular phone data has been used in order to obtain real time information on travel speeds along major roadways. Preliminary results indicate a great potential for this application. The leading existing option of costly installation of speed detectors at dense intervals (~0.5 km) throughout the network makes the WLT alternative even more attractive. Rose (2006) provides a review on cellular phone applications for transportation. He mentions that network-based solutions, which rely on passive monitoring of data already being communicated in the mobile phone system, have the potential to provide network-wide travel time and origin–destination information.

Akin and Sisiopiku (2002) introduced a way to estimate origin–destination matrices with location information from cellular phones. The method is based on mobile phone location data and was tested only by simulations. White and Wells (2002) performed a pilot study in the UK, in the region of Kent, using data from one cellular phone provider. They showed that it is possible to obtain OD information from mobile phone data, but indicated that there is a need to obtain very large samples.

Sohn and Kim (2008) developed a method for the estimation of dynamic OD flows using cell phones as traffic probes. Unlike a direct approach that uses sample OD flows extracted from the cell-based location data as additional observations, and indirect method was proposed wherein the assignment map in the model is derived from the passing time at observation locations and the path choice proportion. The accuracy of the estimation was sensitive to the market penetration. The performance of the estimation was acceptable for proportions of probe phones greater than 5%.

Pan et al. (2006) investigated a cellular-based method to extract trip distribution data for urban transportation planning. A trip distribution table is generated as output with input of spatio-temporal information retrieved from cell phone subscribers. An experiment adopting this method has been implemented in China to test the feasibility of the method. The authors conclude that the experiment exhibit high practicability and effectiveness of the proposed method. The authors argue that for the cellular-based data-extracting method, a sample size of 30% or more should achieve a confidence level higher than 95%.

Caceres et al. (2007) performed a pilot study using a GSM network simulator designed to generate a synthetic database with location registers. The synthetic database was processed mathematically and turned into an origin–destination matrix and traffic counts. Primary results showed great potential of the method. The authors discussed the low precision of the cellular phone data. The precision depends on the radius of the cell, which might be 200 m in a city or between 5 and 20 km in rural areas. Other models providing greater precision do however exist, but they either require modification of the network (installation of ancillary devices to measure the signal, reception angle, signal level and so on) or involve alteration of the terminals (GPS, clocks to record reception times and so on).

Ahas et al. (2007) studied the positioning accuracy when using mobile positioning data in geographical studies. The calculations of theoretical positioning error based on 180,000 positioning measurements in the Estonian GSM network in 2004 showed that 61 percent of positioning points are accurate to within 1000 meters in urban areas and 53 percent are accurate to within 3000 meters in rural areas. They concluded that despite the low accuracy of the positioning data, many research directions, which use a smaller scale such as commuting and regional development studies are nevertheless very promising.

The conclusion from the studies above is that it is possible to use cellular phone data to estimate directly OD trips, provided that the traffic zones are sufficiently large. The study reported in this paper follows the recommendation of the previous studies, focusing on long-distance trips for a national transportation model.

3. METHODOLOGY

In Israel, the use of cellular phone is widespread. Given the small size of the country, cellular phone companies provide national coverage. Figure 1 shows the distribution of the antenna positions for a specific cellular phone provider. As expected, the density of the antennas is higher in urban areas. In less populated places, the antennas are placed close to national roads.

The penetration of cellular phones to the Israeli market, even to lower income households, is quite overwhelming, as shown in Figure 2. It is reasonable to anticipate that cellular phones are even more widely used among individuals in the ages of 10-70, who are the main focus of travel behavior studies. Therefore, we assume that cellular phones data can reasonably represent travel behavior of the entire population.

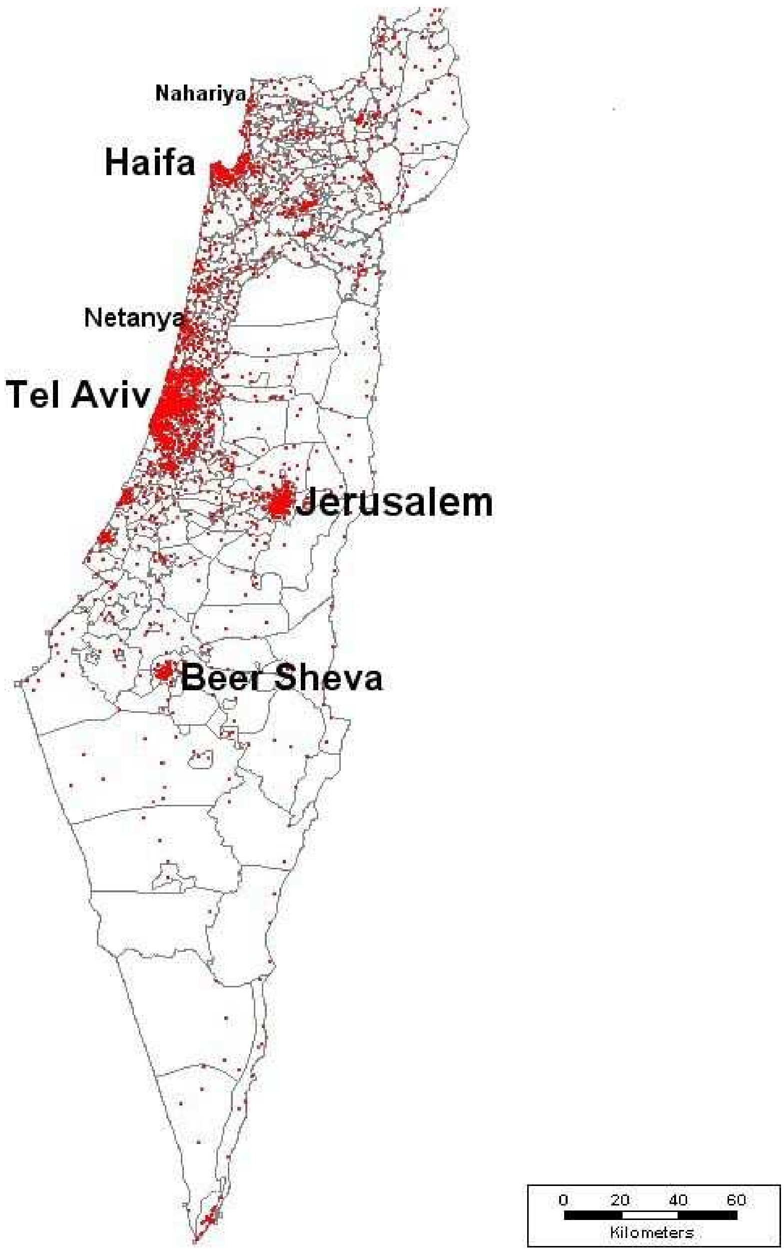

Figure 1. Cellular Antenna locations in Israel, July 2007 (source: Israel Ministry of Communications)

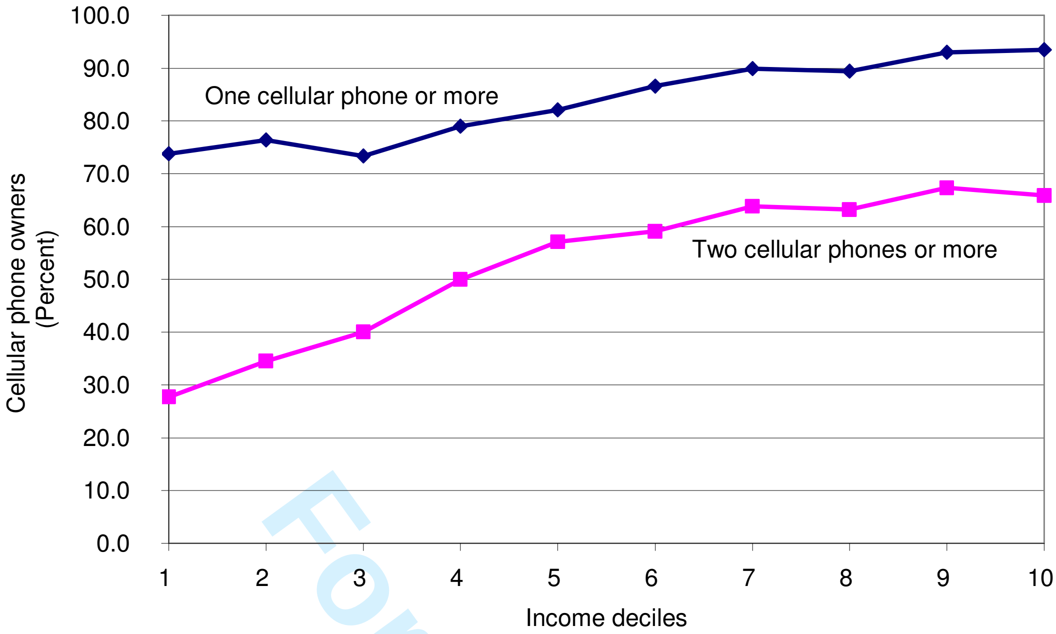

Figure 2: Number of cellular phones owned by household by income deciles, July 2007.

Direct usage of cellular phones data, like the dataset collected in this study, implies several non-trivial assumptions. The main ones are: cellular phones remain open and with their owners; a single cellular phone is carried by the same person; a single person does not carry more than one cellular phone of the same provider at the same time; data from a single cellular service provider (as in our case) is representative of the entire population. Some of these issues have been addressed in this paper, and are discussed in the subsequent subsections.

The methodology used in this exploratory project was fully supported by the Economic and Planning Department of the Israel Ministry of Transport, and is composed by three main steps, which are described in the following subsections. The first subsection summarizes the results of a pilot study conducted prior to the main data collection effort. The second subsection presents the results of a telephone survey conducted to obtain general information about cellular phone habits and provide expansion factors. The third subsection describes the cellular phone data collection, including sample size considerations and data processing.

2.1 Pilot Study

A pilot study was conducted in 2006 to provide an indication if the data can be collected and processed according to the purposes of the study. A sample of 916 cellular phone numbers was randomly selected, and data was recorded for one week. On average, cellular phones were open 69% of the times. As show in Figure 3, about half the phones remained open and provided data to the system more than 90% of the time. Figure 4 shows that for 77% of the phones there is at least one continuous sequence of observation for more than 24 hours during which the phone was open, thus providing data for a duration that is equivalent to travel diary in regular household surveys. The findings suggested that cellular systems could provide significant data on travel behavior.

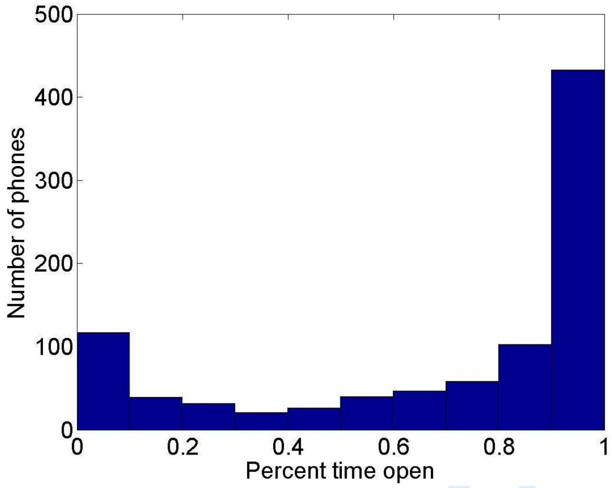

Figure 3: Distribution of 916 cellular phones by percent time they were open.

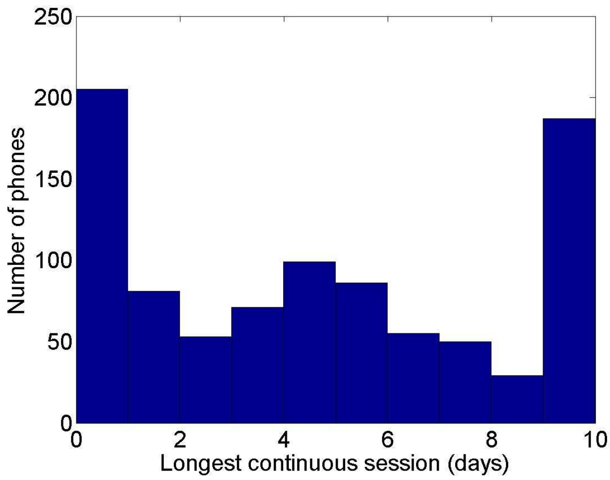

Figure 4: Distribution of 916 cellular phones by the longest continuous period of time (in days) during which they were open.

A reasonable concern is that cellular phone owners are not necessarily a representative sample of the entire population. This is indeed a concern in any type of survey. A first step towards the evaluation of potential biases in cellular phones data can be an inspection of geographic distributions. It seems reasonable to assume that for most people their regular location during the night hours is their home. The night stay location of each phone can be determined from the cellular system data. The distribution of these locations into regions of Israel can be compared with available statistics about the distribution of the entire population into the same regions. This comparison is shown in Table 1.

	Region	**Population**		**Cellular phones night stay**	
		Thousands	**Percents**	**Daily average**	**Percent**
1	Northern Galilee	642	9.3%	47	6.5%
2	Southern Galilee	929	13.5%	76	10.5%
3	Haifa	451	6.5%	49	6.8%
4	Central District – North	1,099	15.9%	129	17.8%
5	Central District – South	946	13.7%	88	12.1%
6	Tel-Aviv	998	14.5%	102	14.1%
7	Judea and Samaria	167	2.4%	50	6.9%
8	Jerusalem	872	12.6%	80	11.0%
9	South District	392	5.7%	58	8.0%
10	Negev	404	5.9%	46	6.3%
	Total	6,900	100.0%	725	100.0%

Table 1: Comparison of distributions of observed cellular phone night stay with population by district.

The table shows that there is a clear correlation between the two distributions, even though they are not perfectly identical. The Pearson correlation between the two distributions is 0.86, and the

Kolmogorov-Smirnov test statistic for two populations is 0.075, meaning that is not possible to reject the hypothesis that the distributions are similar. The match is particularly satisfactory in the three main metropolitan areas, Jerusalem, Tel-Aviv and Haifa. The results of the pilot study show that it is possible to associate weights with observations from the cellular system in order to make the results more representative of the entire population. However, since the data collection is obtained from a single cellular phone provider, it is needed to verify if the sample is representative of the entire population. The next section presents the results of a cellular phone travel habits survey that was conducted in parallel to the main data collection phase.

2.2 Cellular Phone Habits Survey

The collection of the data indicated in the previous section was performed without any contact with the individuals, that is, this was a passive data collection. The results of the pilot survey indicated that there is a need to correct for different market shares in the population of the cellular phone provider. Therefore, a separate specific survey was conducted to complement the passive data collection. This survey was conducted in parallel to the data collection effort, between April and May 2007.

The survey method is similar to household travel surveys, and was performed by a Computer-Aided Telephone Interview (CATI) method. The sample was extracted randomly from the national telephone list. Full details of the survey can be found in a report prepared for the Israel Ministry of Transport (Cohen, 2008), and the results are briefly summarized as follows.

83% of the persons of age over 8 in the sample possess at least one cellular phone. Among ages 18-60 the average rate is 93%. These high rates are observed across several population segments: 95% of workers, 93% of students, and if a person study and work, the rate is 99%. In Israel there are 3 main cellular phone providers. Figure 5 shows selected penetration rate results.

Figure 5. Penetration Rate of Cellular Phone for selected Population Segments.

The market share of "Orange" found in the survey was 30%, but it is not evenly distributed across the population: among young adults (age 18-24) the share is higher (37%), and among Arabs the share is lower (21%).

A specific question in the survey asked about the time that the cellular phones are turned on, and if the person carries the cellular phone during his/her travel. It was found that in 94% of the cases trips can be tracked, since only 6% of the cases the phones are off or left home. The results of this survey were used as expansion factors in the data processing phase. For privacy reasons, the expansion factors were calculated based on geographical regions (defined in Table 1), since no personal information was collected in the cellular phone tracking data described in the subsequent section.

2.3 Cellular Phone Data Collection

2.3.1 Sample Size Considerations

It is well known that household surveys cannot be used to estimate directly origin-destination (OD) trips. Ortúzar et al. (1998) analyzed the number of trips by OD cell in Santiago for a group of 34 zones using data from the 1991 OD survey. They observed that only 58% of the OD cells contained more than 1,000 trips. Thus, according to Smith (1979) it would seem necessary to postulate a sample size of 4% of trips to estimate an OD matrix with a 25% standard error and 90% confidence limits. Assuming 75% response rate, this would imply an initial sample size of 70,000 households.

The cellular phone method allows for high sample sizes with low costs, and because it is a passive data collection method, non-response is alleviated. The main problem is related to persons that travel without cellular phone; however, according to the cellular phone habits survey (presented in the previous section), 94% of the persons regularly take their cellular phones with them. Since the present study focuses on long-distance trips for a national model, it is reasonable to assume an accuracy of 90% for trips longer than 2 km in urban areas and 90% for trips longer than 10 km in rural areas (because of the less dense cellular phone network).

According to historic data in Israel, it is estimated an average of 1,000,000 person trips per hour at the national level (that is, trips longer than 2 km in urban areas and longer than 10 km in rural areas) in peak hours. Given the high coefficient of variation between OD pairs, the only way to obtain a reasonable estimate is to obtain a very large sample. To achieve a high number of observations, the idea was to perform repeated random sampling, to allow for an increase of the sample size for a pre-specified number of cellular phones. The next section details the data collection procedure.

2.3.2 Data Collection

The dataset was obtained from a survey was performed by ITIS Traffic Services Ltd. The cellular phone positions were provided by "Orange", one of the main three cellular phone providers in Israel. The raw data was recorded for 16 consecutive weeks between March 7 and July 2, 2007. Every week, an average sample of 10,200 cellular phone numbers was randomly drawn from one cellular phone provider. The method is based on tracking cellular phone positions. Out of 2,785 hours between the survey duration, the data recorded contains information of 2,451 hours, because of technical problems. The net sample contains information about 102 days, of 1.04 million person-days. This value is 7 times higher compared to the last National Travel Habits Survey (NTHS) performed in Israel in 1996/97.

The tracking system was based on recording events that contain a change in the position of the cellular phone with respect to a given antenna. In Israel there are approximately 2,200 antennas, which are not evenly distributed across the country. In dense population and employment regions, the concentration of the antennas is higher. The data collected contained the following information:

- Unique ID for a given telephone number (which is different from the real number in order to maintain privacy). This ID was given by the telephone provider.
- Coordinates of the antenna that is serving the cellular phone.
- Time stamp (date, hours, minutes, seconds).

The minimum recording frequency was set to 2 hours. This value is related to cellular phone provider policy. This means that in case that no cellular phone movement is observed (and the phone is on), the company scans the phone and records its position every 2 hours. In case the cellular phone changes antenna, the movement is recorded. In addition, when a person speaks in the cellular phone, the recording frequency is very high. As a consequence, the data files were quite big, and to analyze the file several such entries were erased. In overall, 79.2 million phone displacements were recorded, which corresponds to an average of 775,000 displacements per day, or 76 movements per cellular phone.

2.3.3 Data Processing

The large amounts of data collected (79 million positions) was processed in a systematic way described as follows. In the first phase, the data was cleaned according to the following rules:

- Elimination of very frequent observations in which the same ID was in the same antenna location. Only the first and last observation was kept in the data.
- Elimination of frequent observations in which the same ID switched close antenna locations in very short time intervals. It was assumed that such observations are not related to trips, but are related to the way the cellular phone provider handles the phone communications.
- Elimination of unreasonable movements performed in short time periods between antennas located far apart.

After performing the cleaning phase, the next phase consisted of a creation of a weekly "trip diary" for each ID. The "home" was identified by the longest time periods without antenna changes (which occurred mostly at night). Given that the collection was performed in a weekly basis, there were sufficient days to verify the "home" assumption. Special attention was paid to weekends and holidays, in which the "home" assumption might be violated.

A trip start was identified by the first change in an antenna location, and a trip end was identified by the last change in an antenna location, and no movement was observed at least for 20 minutes. The coordinates of the first and last antenna respectively defined the origin and the destination of a trip. These coordinates were geo-coded to 585 traffic analysis zones (TAZ) at the national level.

The data collected and processed was expanded for each phone ID using the "home" position. The expansion factors obtained in the CATI survey were applied using geographical position of the home location. The proportion between the total population in an area (age 8 and over) and the number of phone-days for the same area was calculated to obtain the expansion factors. These expansion factors range between 5 and 20 for weekdays.

3. RESULTS

As discussed in the previous section, the matrix estimation does not include local short trips, because of the accuracy of the system. Since the focus of the research is on long-distance trips, the minimum distance set for a trip was 2.5 km. The system is able to accurately capture trips longer than 10 km. Between these two values, the accuracy is dependent on the frequency of the trips and antenna locations.

The processed cellular phone was compared with the national survey conducted in 1996/7 by the Central Bureau of Statistics for the Ministry of Transport. Although the data collected in this survey is 10 years old, it is still the most recent nationwide household travel survey conducted.

3.1 Trip Rates

The aggregate results found from the whole sample indicate that an average of 12 million person-trips a day on weekdays (Monday-Thursday) was recorded. This corresponds to an average of 2.0 trips per person aged 8 or more. On Fridays the average is 7.7 million person-trips, which is 64% of the average weekday. On Saturdays the average is 6.02 million person-trips, approximately 50% compared to the average weekday. The lower values on Fridays and Saturdays are expected since they are not regular working days: Friday is considered half day for working purposes and several public and private offices are not opened, and religious Jews do not travel on Saturdays.

The average trip rate for weekdays (2.0) is higher in comparison to the 1996/97 NTHS (1.9). However, given that only trips longer than 2.5 km were analyzed, this average is in fact significantly higher compared to the NTHS. This is explained by the fact that the NTHS data is 10 years old, and also because of non-response bias in conventional surveys.

3.2 Trip Length

The average weekly trip length is 17.7 km, but it is not evenly distributed across the week. On Sundays (first working day of the week in Israel) the average is 18.2 km, and the higher value is explained by the movements of students and soldiers. On Fridays the average trip length is 16.6 km, and on Saturdays the highest average trip length is found: 19.4 km, which is explained by trips to tourist places in the north and south of the country. When considering Saturday afternoon trips only (between 15:00 and 17:00), the average trip length rises to 24.1 km. Table 2 shows the trip length distribution for each representative day of the week, and Figure 6 shows the average trip length by period of day.

Distance (km)	Total Daily Trips (Thousands)				Distribution (%)			
	Sunday	Mon-Thu	Friday	Saturday	Sunday	Mon-Thu	Friday	Saturday
<10	6,518	6,583	4,498	3,323	54.3	54. 9	58.3	55.2
10-20	2,413	2,384	1,464	1,162	20.1	19.9	19	19.3
20-30	1,141	1,158	675	543	9.5	9.7	8.8	9
30-40	626	641	350	291	5.2	5.4	4.5	4.8
40-50	364	350	193	158	3	2.9	2.5	2.6
50-60	258	248	143	123	2.1	2.1	1.9	2
60-70	192	182	100	96	1.6	1.5	1.3	1.6
70-80	98	88	55	54	0.8	0.7	0.7	0.9
80-90	81	75	42	42	0.7	0.6	0.5	0.7
90-100	68	63	38	38	0.6	0.5	0.5	0.6
>100	249	222	152	190	2.1	1.9	2	3.2
Total	12,009	11,994	7,711	6,021	100	100	100	100

Table 2. Trip Length Distribution (Trips longer than 2.5 km)

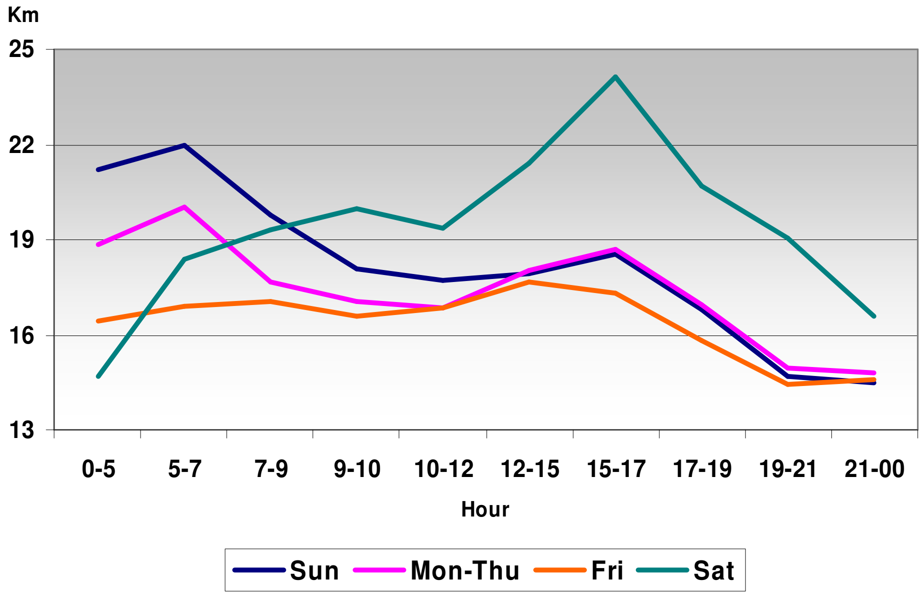

Figure 6. Average Trip Length by Period of Day

Trip rates calculated from cellular phone positions for trips with average distance greater than 25 km are significantly higher compared to the NTHS (38% more). The NTHS recorded trips for 3 consecutive days, and only 1% of the households were sampled. The short number of surveyed days and small sample is not enough to capture long trips.

3.3 Origin-Destination Tables

The data collected enabled the estimation of origin-destination (OD) tables for 10 different periods of day and 4 representative days. In addition, special OD tables for holiday periods were built, to assess the movements in special occasions. For the first time such comprehensive tables were built at the national level.

Figure 7 compare the results between the cellular phone data and NTHS data for selected OD pairs. The average population increase between 1996/97 and 2007 is estimated at 25%, therefore it is expected that the cellular phone data will produce more trips compared to the NTHS. However, given that the cellular phone data is 72% higher compared to the NTHS, the difference cannot be explained only by the population and motorization rate increase. The cellular phone data covers the entire population, while the NTHS survey is accurate at "regular" households. Special population groups such as students, soldiers, foreign workers and so on were under-represented in the NTHS survey.

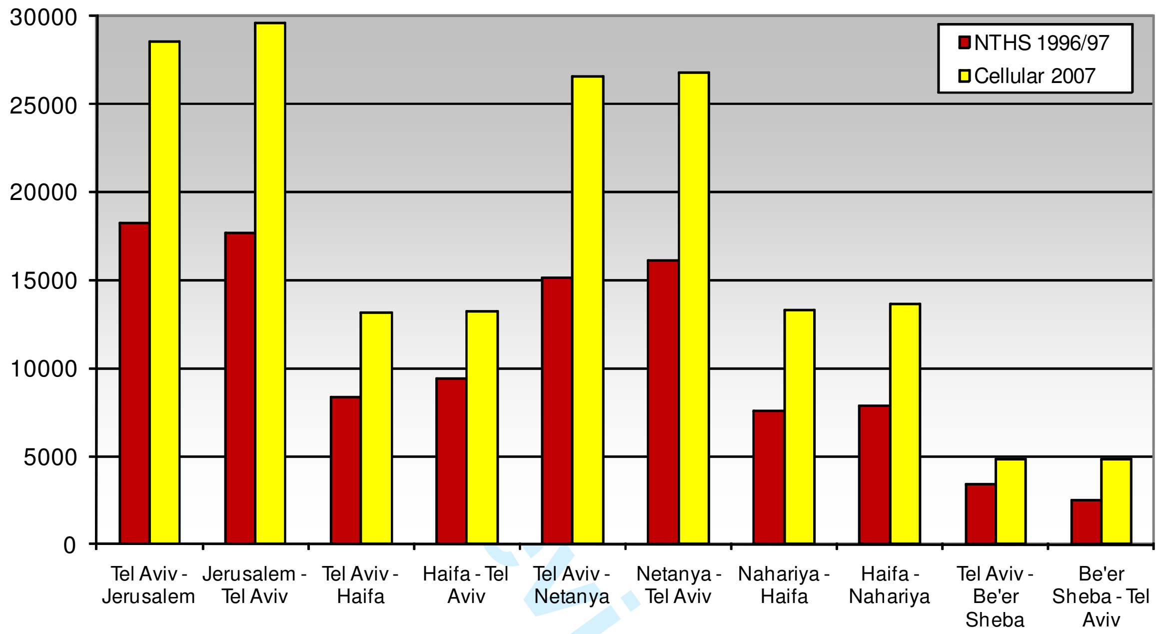

Figure 7. Comparison between NTHS and Cellular phone Trips for Selected Origin-Destination Regions.

4. SUMMARY AND CONCLUSIONS

Typical household surveys collect data related to a limited number of days for a relatively small sample, and therefore less frequent trips are under-represented. This paper presented a methodology to track cellular phone positions at a national scale. The method allows the collection of large amounts of data at relatively low cost, and overcomes a well-known deficiency of conventional surveys, which is the under-representation of long distance trips. These advantages enable monitoring travel demand at the aggregate level, and could be used in several transportation and land use applications. A recent paper by Gur et al. (2009) shows the application of the data collected in this study for calibration of aggregate trip generation and trip distribution models.

Given the nationwide antenna coverage, the method enables simultaneous data collection at different places, thus enforcing consistency. The study allows the analysis of person-trips not only for weekdays and peak periods, but also for other days of the week and different time periods. New transportation planning models are activity-based, and therefore it is important to obtain a complete description of the trip pattern. The method allows the validation of transportation planning models at the OD level.

The transformation from antenna positions to person-trips required several assumptions. Further investigation requires special surveys in order to estimate the extent by which these assumptions are violated, the resulting biases and the possibilities to correct for these biases. Such bias evaluation is not part of the scope of this research, in which we assume that the biases are small and correctable and focus on evaluating the potential of the method by examining the available cellular phone data.

There are several limitations related to data accuracy and resolution. The method presented in this paper used the closest antenna position and related it to a traffic analysis zone. Inspection on the data revealed many "strange" patterns, such as zigzagging between adjacent zones. Further research is needed to validate the accuracy of the method and refinement to smaller zones.

Another limitation of anonymous cellular phones data is that it does not include direct information about travel purpose and traveler characteristics. These characteristics are known as key factors in travel behavior studies. Equivalent information can be implied from the raw data itself, but only to a certain extent. Further research is needed to assess the quality of data that can be inferred from this passive method.

It should be noted that the proposed method of collecting OD data is not intended to replace household surveys, but to complement household survey data. Because of privacy, no personal information is collected, and therefore no individual behavior can be modeled with the method. In addition, the data obtained from such method is only limited to trip timing, length and duration. It is difficult to identify trip chaining; Gur et al. (2009) presented a method to capture simple and compound tours, but unless accuracy of the method is improved, it is still a challenge to form a daily tour pattern. Since the method is based on cellular phone displacements, it is also not possible to identify the number of people who are telecommuting; or working at night (because of the assumption of the "home" place). To overcome such problems, complementary surveys must be conducted.

It is noticeable that the active participation of the Israel Ministry of Transport enabled the completion of this unique project at a reasonable amount of time. A project of this nature requires the direction of a centralized agency dealing with transportation planning and analysis, in order to overcome difficulties that may arise in the several tasks involved in the project.

The OD matrices derived from the cellular phone data are currently being used in other projects by the Israel Ministry of Transport. Traffic assignment models were performed using these matrices, and initial reports show a satisfactory match with conventional traffic counts. Demand modeling based on the cellular phone data is reported in a separate paper (Gur et al., 2009).

REFERENCES

Ahas R., Laineste J., Aasa A. and Mark U. (2007). The Spatial Accuracy of Mobile Positioning: Some experiences with Geographical Studies in Estonia, In Location Based Services and TeleCartography, G. Gartner, W. Cartwright and M.P. Peterson (Eds.), Springer, pp. 445-460.

Akin, D. and Sisiopiku, V.P. (2002). Estimating origin–destination matrices using location information from cell phones, Proceedings of the 49th Annual North American Meetings of the Regional Science Association, Puerto Rico.

Axhausen, K., Zimmermann A., Schönfelder S., Rindsfüser G., and T. Haupt. (2004). Observing the rhythms of daily life: A six-week travel diary, Transportation, 29(2), pp. 95-124.

Caceres, N., Wideberg, J.P., Benitez, F.G. (2007). Deriving origin–destination data from a mobile phone network. IET Intelligent Transportation Systems, 1(1), pp. 15–26.

Cohen, Y. (2008). Evaluation of travel patterns data from cellular systems (in Hebrew), Intermediate report to the Israeli Ministry of Transport.

Gur Y., Bekhor S., Solomon C., and Kheifits L. (2009). An Aggregate National Transportation Planning Process in Israel: Formulation and Development, Transportation Research Record: Journal of the Transportation Research Board, 2121, pp. 145-151.

Pan, C., Lu J., Di S., and Ran B. (2006). Cellular-Based Data-Extracting Method for Trip Distribution, Transportation Research Record: Journal of the Transportation Research Board, 1945, pp. 33-39.

Qiu, Z., and P. Cheng (2007). State of the Art and Practice: Cellular Probe Technology Applied in Advanced Traveler Information System. Proceedings of the 86th Annual Meeting of the Transportation Research Board (CD-ROM).

Richardson, A.J. and J. Wolf. (2001). Data Structures, Sampling and Survey Issues: Report of Workshop M6. 9th International Association of Travel Behaviour Research Conference, Gold Coast, Australia.

Rose, G. (2006). Mobile Phones as Traffic Probes: Practices, Prospects and Issues, Transport Reviews, 26(3), pp. 275–291.

Sohn, K. and D. Kim (2008). Dynamic Origin–Destination Flow Estimation Using Cellular Communication System. IEEE Transactions on Vehicular Technology, 57(5), pp. 2703-2713.

Stopher, P.R, Wilmot, C.G., Stecher, C.C., and R. Alsnih (2006). Household Travel Surveys: Proposed Standards and Guidelines, In Travel Survey Methods: Quality and Future Directions, P. Stopher and C. Stecher (Eds.), Elsevier, Oxford, pp.19-74.

White, J. and Wells, I. (2002). Extracting origin destination information from mobile phone data, 11th International Conference on Road Transport Information and Control (Conf. Publ. No. 486), London, pp. 30–34.