Intercity Person Trip Table for Nationwide Transportation Planning in Israel Obtained from Massive Cell Phone Data

Yehuda J. Gur - Engineering Consultant

Shlomo Bekhor - Associate Professor, Transportation Research Institute, Technion (* corresponding author)

Charles Solomon - Deputy Director General, Economics and Planning Department, Ministry of Transport and Road Safety

Leonid Kheifits - Planning Consultant

המאמר התפרסם ב This article was published in TRR journal no.2121 (2009), p. 145-151 (Transportation Research Record)

Intercity Person Trip Table for Nationwide Transportation Planning in Israel Obtained from Massive Cell Phone Data

ABSTRACT

A need for a national transportation planning model has risen for a number of reasons: heavy investments in transportation infra-structures, and the need to perform formal cost-benefit analysis of all medium and large scale transport projects. The paper describes the formulation and development of the planning model.

The model is unique in terms of the data collected and used for analyzing nationwide travel. After careful legal review regarding privacy laws, cellular phone data were obtained for 16 one-week samples of 10,000 phones. In total, data for 1.04 million person days were obtained. Data records included: the unique cell-phone ID, the antenna serving the cell-phone, and a time stamp (date, hour, minute, second). At the minimum, a record was written each time a moving cellular phone changed its connecting antenna. To ensure privacy, neither information nor identification of the cellular phone owner was recorded.

The article describes the structure of the planning model. The cellular phone survey data were used only in the models for constructing person trip tables. To our knowledge, this is the first time where data obtained by Wireless Location Technology (WLT) in large quantities are used in transportation planning. The paper discusses the advantages and limitations of using WLT and presents directions for further research.

INTRODUCTION

In recent years there has been a quick and substantial increase of investment in the inter-city transport network in Israel. Major efforts include:

- Fast development of a countrywide railroad network
- Heavy investment in joint public-private investments in major transport facilities, primarily toll roads in addition to urban LRT's
- Heavy investment in improving the intercity and rural highway network

In parallel, the government decided to implement and apply a formal procedure for the evaluation of the worthiness of transport projects [1]. No large transport project in Israel is financed without proving its worthiness. The evaluation depends heavily on estimates of future travel demand.

The Metropolitan transport planning agencies in Israel face difficulties in estimating the amount and characteristics of external and thru travel. Significant differences between agencies have been found in estimates of the amount of travel for identical interchanges, and they are documented in technical reports [2].

For all those reasons, a pressing need for improved planning tools for non-urban projects was obvious. A decision was made to formulate and establish a uniform country-wide transportation planning process, including a travel demand model. The process had to be tailored to the needs it supposed to satisfy.

First, it was clear that the major usage of the process was to address large scale problems: Long range travel demand predictions, including its dependence on government policies and future scenarios and the analysis of the effects of large scale projects on network performance, etc.

This meant that the basic demand estimates (constructing the person trip table) should produce reliable estimates, but can be of limited refinement and accuracy. More sensitivity was needed primarily in the modal split model which estimated modal shifts. The process should be open; it should have a close two-way interaction with the metropolitan planning processes: estimates of intra-metropolitan travel demands and network performance should be obtained from the detailed metropolitan studies. The metropolitan studies would obtain estimates of external and thru travel from the national model.

Considerations in Selecting the Planning Model

A major drawback of the effort of formulating the planning process was the lack of understanding intercity travel. Intercity travel in Israel was different from most other countries, primarily because of the small country size and the absence of surface travel across the borders. Hence, it was impossible to learn on intercity travel from other countries.

Intercity travel is principally different from urban travel. This is because of the much less pronounced role of regular, commuting travel. We hypothesize that the observed regularity in intercity traffic patterns is primarily a result of the law of big numbers, rather than the regular behavior of individual travelers. Hence, at the outset, it seems less important to use detailed behavioral travel demand analysis, in particular in estimating the person trip table. The approach used in this paper is similar to statewide transportation demand models in the U.S. [3].

The last large scale country-wide travel habit survey in Israel was conducted in 1996 [4]. It did not provide satisfactory information on intercity travel: First, the survey reported only a very small number of intercity trips, not sufficient to enable understanding or model formulation.

Second, the expanded data failed to replicate satisfactorily traffic counts along major cordon lines. Severe underreporting of this type of travel is suspected. The option of performing another travel habit data collection was rejected because of the cost and time involved, and the inherent difficulties of obtaining good enough information. A number of recent studies pointed out difficult problems in conducting such surveys, either by direct contact or by CATI (Computer aided telephone interviews) [5]. The two most serious problems are low response rate, reported to have declined from about 50% in the 60's and 70's to less than 20% in recent studies, and the need for active participation by the respondents, relying on their cooperation, memory and patience [6].

It was decided to use modern data sources which did not require active respondent participation, but rely on Wireless Location Technology (WLT), such as GPS or cellular phones (CP). There were a number of reasons for this selection. It provided large amounts of relevant data at reasonable time and cost. Second, the travel information could be renewed periodically at a reasonable effort, hence there can be continuous tracking of the travel habits and the model can be updated at a reasonable effort.

Statistical analysis of WLT data from cellular phones has been recently used in transportation studies. Qiu and Cheng [7] review twenty studies world-wide were cellular phones 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.

In a recent study, Candia et al. [8] used a similar method to investigate collective behavior at large scales and focus on the occurrence of anomalous events. The authors focused on aspects of human dynamics and social interactions. The paper does not provide details on the sample characteristics, but the conclusions clearly indicate that large-scale collective behavior can be described using aggregated data resolved in both time and space.

CELLULAR TELEPHONE SURVEY

CP Data Collection

The major data source utilized location data from a CP system. The penetration of CP's in Israel, even to young children and low income families is quite overwhelming. Even in the lowest income decile, 75% of the households possess at least one CP. In median income households, 82% possess at least one CP, while 67% possess more than one CP. It is reasonable to assume that most of the mobile individuals, in the ages of 10 to 70 which are the main travel makers, possess CP's cellular phones [5].

The CP survey was a passive data collection; that is, it did not require any contact with the individuals. The data were provided by the ITIS, Traffic Services Ltd. ITIS obtained the data from the "Orange" CP Company, one of the three major CP providers in Israel which operates close to 2 million phones.

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 was determined from the cellular system data. The distribution of these locations into regions of Israel was compared with available statistics about the distribution of the entire population into the same regions. This comparison is shown in Table 1.

[Insert Table 1 here]

The table shows that there is a clear correlation between the two distributions, even though they are not perfectly identical. The match is particularly satisfactory in the three main metropolitan areas, Jerusalem, Tel-Aviv and Haifa. Clearly it is possible to associate weights with observations from the cellular system in order to make the results more representative of the entire population.

The "Orange" company covers the country with approximately 2,200 antennas. Each (operating) CP is connected to the closest antenna. A moving phone changes periodically the antenna. Every such a change is recorded as a part of the standard operation. For our particular needs the data only had to be accumulated and slightly processed. The data collected contained the following information:

- Unique ID for a given CP number (which is different from the real number in order to maintain privacy). This ID was given by the CP provider.
- The ID of the antenna that is serving the CP.
- 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 the CP is idle (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 it was necessary to erase several such entries prior to the analysis. .

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 sample was drawn with replacement: any CP could not participate in more than one sample

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.

CP Habits Survey

There was a need to correct the survey for possible bias in the distribution of active "Orange" CP's in the population. Therefore, a separate 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 was a conventional Computer-Aided Telephone Interview (CATI) survey [9]. A sample was extracted randomly from the national telephone list. The information obtained from the survey included the geographical distribution of owners of CP's and the behavior of phone owners with regard to keeping the phone turned on and carrying the phones while traveling.

According to the survey results, 83% of the persons of age over 8 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. 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%).

The survey asked specifically 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 the survey clearly indicated biases in the distribution of active "Orange" CP's. These results were used to calculate expansion factors to weigh the raw CP data.

DATA ANALYSIS

For planning purposes, the country was divided into 600 traffic analysis zones (TAZ). Each of the 2200 antennas was assigned to a TAZ. This TAZ was considered as the location of the phone. This process was not accurate; the selection of an antenna by the system in any given moment is dependent on many factors, such as quality of connection, the direction of the phone, etc. The assumption that the phone is located in the same TAZ as the antenna is inaccurate. This problem affected significantly a number of elements in the model formulation and calibration. First, it was decided to ignore short trips, which their description in the data would have been highly erroneous. This caused certain incompatibility in the definition of trips between the countrywide model and the classical demand models. The loss of short trips was expected to have only small effect on the estimated traffic volumes of major facilities. Second, many trip ends were located not in their correct zone, but in a neighboring zone. This had a harmful effect on the trip generation model, which related the number of trip ends in a zone to external data on activities' types and levels.

Processing of the survey data included the following steps:

- Removing redundant and erroneous records
- Processing the weekly diary of each CP:
  - Definition of a trip end as a stay longer than 20 minutes in the same TAZ.
  - For each trip end: arrival, departure and dwell time (intra-TAZ trips were not counted)
  - Identification of "Home": the TAZ with the longest total weekly dwell time
  - Aggregation of consecutive 'trips' which were within 2 km radius
  - Definition of tours: a chain of trips which started and ended at "home". For each tour, main destination TAZ (The longest dwell time) was identified
  - Removing short, 'irrelevant' tours: maximum distance is less than 4 km, or total tour time is less than 20 minutes, or no stop along the tour is longer than 15 minutes
  - Chaining sections of tours with dwell time of less than 15 minutes
  - Assigning each trip (tour leg) with trip type code and period of day (POD) code, as appears in Table 2.

[Insert Table 2 here]

The selection of the way to classify the trips required some analysis. In standard planning processes, trips (or tours) are classified first by trip purpose and models are formulated specifically for each trip purpose. However, this data item was not available in the CP survey. The only variables available for classification were: the "home" trip end and the tour's passively obtained attributes.

Classification of Tours and Trips

Five trip types were defined: for simple (2-leg tours), the first leg corresponds to the trip from home, and the second leg is the trip to home. For compounded (3 or more leg tours), the first leg is the trip from home, the last leg is the trip to home, and the remaining trips from the tour are non-home based trips.

As seen in Table 2, almost 75% of the tours are simple, one-way tours. The different trip types are indeed different: The average total length of compounded tours is significantly longer than that of the simple trips. On the other hand, the legs of the simple tours are significantly longer than the other legs: 21.55 km vs. 7.6 – 9.4 km in the compounded tours. This implies that the trip distribution models of the different trip types will have to be different. The total length of the compounded tours is about 40% longer than that of the simple tours.

For modeling purposes, 8 periods of day (POD) were defined: the 24 hour weekday was subdivided into 7 periods of duration of 3 to 5 hours. In addition, the weekend was defined as one more POD. . This was necessary in order to model tours which start in the weekend but terminate during the week, or vice versa.

Within the model, the daily number of trips was defined as the total number of trips starting during weekdays of a typical week divided by 5, plus the total number of trips starting during a typical weekend.

Characteristics of the Survey Data

At the outset it was clear that the data had some limitations. First, it did not provide combined information about travelers and travel. Hence, it did not permit the formulation of models which related travel making to travelers' characteristics. Second, it was found that the available data had accuracy limitations: Trip ends' location was determined with limited accuracy; in many cases, a trip end was located not in the zone where it actually happened, but in a neighboring zone. Hence straight regression models, which relate the number of trip ends to zonal characteristics, were necessarily of limited accuracy.

The "home' location were defined as the longest stay TAZ. This definition was not fully compatible with the way the zonal population is estimated, in particular for intern population, such as soldiers, students, etc.

The data were corrected to account for the distribution of active "Orange" CP's in the population. Still the data described the movement of CP's, rather than of people. In order to examine the seriousness of this incompatibility it was decided to compare the survey data to an extensive set of traffic counts, many arranged in cordon and screen lines, and if necessary, to correct the survey data accordingly. Happily, a very good match was found between the two data sets; no correction was necessary.

A serious drawback of the data was the lack of any information on the mode of travel. For this, as well as other reasons, the CP survey data was used only for the estimation of person trip tables. The mode split model had to be based on other data.

The major advantage of the data was the detailed and complete picture of country-wide travel it provided. The picture was not accurate on the level of 600 TAZ, but seemed to be very satisfactory on level of 34 districts. They provide a direct picture of the connection between the trip table and the traffic counts.

The detailed, survey-based trip tables provided a strong base for a person trip table construction.

THE TRAVEL DEMAND MODEL

Figure 1 describes the structure of the travel demand estimation process. It is a slightly modified version of the standard four-step process adjusted to the particular needs and characteristics of the problem in hand.

The process relies heavily on the assumption that the present day information on travel characteristics (from the CP survey) is of a good quality and should be preserved as much as possible. Hence, the process is based on a "change" model, which estimates changes in the present (survey based) travel demand because of possible external changes: future activity scenarios, new transport facilities, new policies, etc. The process is designed to operate closely with the planning processes of the metropolitan planning agencies: It is assumed that these processes produce better estimates of travel demands and network loads and travel times inside their areas; hence, it makes sense to obtain their estimates of these data.

The purpose of the process is to estimate the changes in travel demand due to external changes, in a target year. In its major application the model estimates traffic loads for 3 typical periods of day for an average weekday (in Israel, weekdays are Sunday thru Thursday; Friday and Saturday are weekend).

The model deals with three planning scenarios: "The base year", 2007, where the survey has been performed, "the base planning year", the most likely scenario of a future year (say 2030), for which predicted activity levels and network are known, and "the examined scenario", the base planning year with changes whose effects are examined. In order to insure maximum accuracy in the analysis, the description of the examined scenario uses, as a base for changes, the base future scenario.

The input to the process includes the following elements:

- Multi-modal networks and TAZ activity data for the base target year
- Base year data, including TAZ activity and person trip tables
- Data from the metro planning agencies, including travel times on the arterial network
- Specifications of the differences of the present scenario from the base target year.

Output of the process includes estimates of traffic volumes and travel times on all network facilities. In addition, external and thru trip tables can be provided to the metropolitan planning agencies.

[Insert Figure 1 here]

The process includes the following major models:

1. Adaptation of the base planning year data to the characteristics of the scenario being examined
2. Car availability model. It should be noted that in Israel, there are many employer owned automobiles driven by employees. Hence, it is necessary to analyze automobile availability rather than ownership.
3. Person trip table construction. The CP survey data was used in this step, which includes two models:
    a. Trip generation
    b. Trip Table adaptation (rather than trip distribution)
4. Modal split
5. Trip table correction
6. Traffic assignment (Highway and public transport)

PERSON TRIP TABLE ESTIMATION

A major component of the process is the trip table estimation. The process is based on a number of principles:

- It has been found that the survey's district level (34 zones) trip table is of a good quality. It is also assumed that the basic structure of the countrywide trip table will not change much in time. Hence, it has been decided to base the trip table estimation on correcting the survey's district table to the trip ends estimates, while changing it as little as possible.

- In the survey trip table, the intra-district distribution of trip ends among TAZ is not satisfactory. Hence, the number of zonal trip ends is estimated using a trip generation model and ignoring the survey data.
- Countrywide travel includes many trips with long stay at the destination, many of which are multi-day trips. Hence, there is a need to model carefully the temporal distribution of the return trips to home.

Trip Generation Model

The trip generation models were estimated using regression analysis. They were calibrated on the TAZ data. The production model estimates the total daily trips departing from home. The functional form is given as follows (eq.1):

$$P_i = a + b[POP8]_i\left(1 + c[CAR]_i\right) \qquad (1)$$

Where:

P_i : The number of weekday tours from home produced in zone i

$POP8_i$: Population (of age 8 plus) of zone i

CAR_i : Number of vehicles available to residents of zone i

a, b , c: parameters

An effort was made to calibrate the model with the parameter a=0, but the estimation errors were too big. The data did not permit the formulation of more complicated models. Even at this level the estimation errors were quite large. Apparently, the data limitations resulted from the inaccuracies in determining the location of trip ends.

The attraction model estimates the number of daily person-tour destinations in each TAZ. It is also estimated using regression analysis and is of the form (eq. 2):

$$A_i = d + e[POP8]_i + f[JOB]_i \qquad (2)$$

Where:

A_i : The number of weekday tours attracted to zone i

$POP8_i$: Population (of age 8 plus) of zone i

JOB_i : Working places in zone i

d, e , f: parameters

The model is calibrated separately for each trip type. Stratification of the employment data by type (e.g., high education, industrial centers, CBD's, etc.) did not contribute to the model's performance.

Within each trip type, the model is calibrated separately for each POD. This approach was selected for a number of reasons. First, there were a significant number of multi-days tours. The selected approach permits a correct description of the temporal structure of such tours. Second, a major phenomenon which the model was supposed to describe was the accumulation of traffic in city entrances during the peak periods. Trips from peripheral zones which contribute to congestion start earlier than the peak period. The model was designed to permit, in the future, the construction of trip tables for dynamic assignment, which is capable of describing this phenomenon.

Trip Table Adaptation Model

The purpose of the model is to translate the base, district trip table to the planning year estimates of TAZ trip ends.

The model is applied separately for the each trip types and for each separate POD's. This is done because of the large difference in the trip tables between the different trip types and POD's.

The model operates as a 2-level iterative proportional fitting (IPF) model. At the first level, the model is applied at the level of 34 districts. Trip ends in the zones within each district are summed, and the base district trip table is adjusted to these sums. In the second level, a 3 variable IPF is applied at the TAZ level: Productions, attractions, and district OD's. The District ODs are the result of the first stage.

The structure of this model was selected for a number of reasons. First, the district trip table was found to be of high quality. The process maximizes the preservation of its structure. Second, because of the limited accuracy of location determination, the survey's distribution of trip ends among zones within each district seemed to be unreliable. The distribution of trip ends within the district would be based on TAZ activity data. A third reason for the selected approach was the observed long run stability, or slow growth, in rural link volumes. It seemed logical to assume that this stability indicates a similar stability in travel patterns, which the IPF model preserves.

At this point in the process, the tours' return trips were modeled. The survey data was used to determine the distribution of the POD of the start of the return trip, for each POD of the start of a tour. Distributions were determined for each trip type and distance ring. Sum of multiplications of the trip table by these distributions produced the return trip table.

Trip tables for the 3 planning periods were derived from the detailed POD tables. Those included: average AM peak, average PM peak, and mid-day.

Mode Split Model

A standard existing probabilistic mode split model is applied to determine the number of vehicle and transit trip tables [10]. The model is designed to address specifically inter-urban travel. Three modes of travel are modeled: automobile, rail and bus. The model is sensitive to a number of factors as follows:

- Automobile availability
- Travel impedance to the closest terminal (bus and rail)
- Public transport door to door travel time
- Automobile travel time
- Travel time reliability (typical standard deviation for each mode by time of day
- Total trip cost
- Trip's comfort: The number of transfers and total trip duration.

The model accounts for the fact that trips are tours. The mode split is estimated considering the return trips travel impedance.

The model includes a built-in iterative procedure to balance the service frequencies with the number of passengers.

Highway and Transit Assignments

The mode split model produces vehicle trip tables and transit passenger trip tables. Combined trips, such as "Park and ride" trips, are decomposed using the shortest path between the origin zone and the closest transit station. Standard deterministic assignment models were used to estimate traffic and passenger volumes.

The quality of the survey's results was examined by assigning the survey trips on the coded network and comparing the results to traffic counts. An excellent correspondence was found, a fact that increased the survey's reliability.

In addition to the standard assignment outputs (traffic volumes, etc.), the procedure retains External and thru trip tables for each metropolitan area, which are provided to the metropolitan planning agencies.

SUMMARY AND CONCLUSIONS

The project is currently in the model validation stage, and initial results show a very good match between traffic volumes obtained from the modeling process and traffic counts. The main feature of the planning process is the fact that the CP survey can provide a robust and complete person travel demand pattern at the district level. Tools were developed to estimate travel demand changes due to a variety of demographic, policy, and network changes.

The main innovation in this process, compared to standard statewide models, is fact that for the first time data obtained by tracking cellular phone positions was applied in a transportation planning model. Other information technology means at these dimensions are used in transport planning. Previous studies using cellular phones were limited to tracing individuals who carry an instrument, but their identity is known [11]. This type of application adds a very important tool to the set of tools available to transportation planners.

An important feature of the method is based on the fact that the CP data is continuously accumulated, and the process of obtaining and processing them has been established. It will be quite straight forwards to apply the process periodically in order to update the data base and trace continuously countrywide travel demand. Of particular importance is the fact that the process was legally approved, accounting for privacy laws.

A framework was established for close interaction between the Metropolitan planning agencies and the countrywide process. This framework included corresponding definitions of traffic zones and network structure, and procedures for data exchange. Such interaction will assist both bodies to use in the planning process the best available data.

A number of disadvantages in the process were found. First, it seems that in the survey, the allocation of CPs can be improved. Improvements might include, for example: Distributing the trip end among a number of zones, according to the probabilities of being there. Another possibility is to include in the allocations information about the TAZ activities. Improved procedures for this step in the process will undoubtedly improve significantly the data quality.

Limited experiments with the highway assignment model indicate that it underestimates the severity of congestion in cities' entrances. A possible reason is the fact that in applying the model, peak hour trips are selected according to trips start time, which can be more than one or two hours earlier than the time of arrival to the destination city entrance. A possible way to correct for this fault is to use dynamic assignment. It should be noted that the CP survey provides very good information on the distribution of trips' start ad arrival times.

The CP survey data is proved to be a very reliable source of information about travel patterns. It is safe to assume that these data can be used in may other applications in the transportation planning process. Promising applications can be, for example, in conjunction with travel habits surveys, both in order to correct the survey's expansion weights for under-reporting, or in order to update periodically old survey data.

REFERENCES

1. Shiftan, Y., Sharaby, N., and C. Solomon. Transport Project Appraisal in Israel. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2079,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 136-145.

2. Perlstein-Galit Company Ltd. *Travel Demand Forecasting Model*. Report prepared for NTA – Mass Transit System for the Tel Aviv Metropolitan Area, 2001.

3. Horowitz, A. Guidebook on Statewide Travel Forecasting. Publication FHWA-HEP-99-007. FHWA, 1999.

4. Central Bureau of Statistics. National Travel Habits Survey 1996/7, (in Hebrew). Report prepared for the Israeli Ministry of Transport, 1998.

5. Stopher, P.R. and H.M.A Metcalf. Methods for Household Travel Surveys. NCHRP Synthesis of Highway Practice 236, Washington, DC, Transportation Research Board, 1996.

6. Grove R.M., and M. P. Couper. Nonresponse in Household Interview Surveys. John Wiley & Sons. Wiley Series in Probability and Statistics Survey Methodology Section, 1998.

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

8. Candia, J. Gonzalez, M.C., Wang, P., Schoenharl, T., Madey, C., and A.-L. Barabasi. Uncovering individual and collective human dynamics from mobile phone records. Journal of Physics A: Mathematical and Theoretical, Volume 41, Issue 22, pp. 224015, 2008.

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

10. Cohen, Y. The DIMOS Modal Split model (in Hebrew). Report to the Israeli Ministry of Transport, 1999.

11. Fontaine, M.D., and B. L. Smith. Probe-Based Traffic Monitoring Systems with Wireless Location Technology: An Investigation of the Relationship between System Design and Effectiveness. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1925,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 3-11.

TABLE 1: Distributions of Observed CP Night Stay and Population by Region

Region	**Population**	**Cellular phones night stay**
Northern Galilee	9.3%	6.5%
Southern Galilee	13.5%	10.5%
Haifa	6.5%	6.8%
Central District – North	15.9%	17.8%
Central District – South	13.7%	12.1%
Tel-Aviv	14.5%	14.1%
Judea and Samaria	2.4%	6.9%
Jerusalem	12.6%	11.0%
South District	5.7%	8.0%
Negev	5.9%	6.3%
Total	100.0%	100.0%

Table 2: Trip and Tour Types

Tour Type	Trip Type	Total Trips	Total Trip Length (km)	Percentage of Total Trips	Percentage of Total Trip Length	Average Trip Length (km)
2 Leg tour	From Home	225,299	4,855,122	26.5%	36.6%	21.55
	To Home	225,299	4,855,122	26.5%	36.6%	21.55
Multi-leg tour	From Home	58,725	481,372	6.9%	3.5%	8.20
	To Home	58,725	466,431	6.9%	3.5%	7.94
	Non home-based	281,214	2,639,243	33.1%	19.9%	9.39
	Total Trips	849,262	1,3297,290	100.0%	100.0%	15.63

Figure 1: The structure of the Planning Process