Inferring multimodal itineraries from rich smartphone data

David Montoya
Cofely Ineo, GDF Suez & INRIA Saclay
david.montoya@cofelyineo-gdfsuez.com
david.montoya@inria.fr

Serge Abiteboul
INRIA Saclay & ENS Cachan
serge.abiteboul@inria.fr

ABSTRACT

We designed a system to infer the multimodal itineraries traveled by a user from a combination of smartphone sensor data (e.g., GPS, Wi-Fi, inertial sensors), personal information, and knowledge of the transport network topology (e.g., maps, transportation timetables). The system operates with a *Multimodal Transport Network* that captures the set of *admissible multimodal itineraries*, i.e., paths of this network with weights providing the statistics (expected time and variance) of the paths. The network takes into account public transportation schedules. Our Multimodal Transport Network is constructed from publicly available transport data of Paris and its neighbourhoods published by different transport agencies and map organizations. The system models sensor uncertainty with probabilities, and the likelihood that a *multimodal itinerary* was taken by the user is captured in a Dynamic Bayesian Network. For this demonstration, we captured data from users travelling over the Paris region who were asked to record data for different trips via an Android application. After uploading their data into our system, a set of most likely itineraries is computed for each trip. For each trip, the system displays recognized multimodal itineraries and their estimated likelihood over an interactive map.

Introduction

The democratisation of connected and sensor-rich personal mobile devices has increased the demand for context-aware commercial applications taking advantage of the information they are able to generate. Most prominent examples of such applications are the smartphones' navigational applications. Modern smartphones can track the position of their carrier using three independent radio networks: Satellite Navigaton (GPS, GLONASS), Cellular and Wi-Fi networks. Embedded inertial sensors (accelerometer and gyroscope) can be used to enrich positional information via *dead reckoning*. And more recently, inertial sensors have also been successfully used in different *activity recognition* applications. Current technology is far from taking full advantage of positional and inertial data to understand the users' daily movements in urban environments. The present system demonstrates sophisticated techniques towards this goal.

We designed a system to infer the multimodal itineraries traveled by a user from a combination of smartphone sensor data, personal information and knowledge of the transport network topology. The sensor data is recorded by users over the course of a journey. The network topology can be constructed from available geographic data and public transport timetables. The system ultimately aims at two modes of operations, each with distinct goals:

Offline mode After the data has been acquired, the system determines the user's location, mode of transport and different transportation routes and lines taken over the time of a travel. The expected result is a set of candidate itineraries ranked by their likeliness.

Online mode In real time, the system determines the current user's location, mode of transportation and current transportation route if applicable. It predicts the user's future movements, both in the short term (Is the user going to grab a bicycle at the next bicycle-sharing station?) and in terms of distant goals (Is the user going to the office?).

This demonstration will most likely only feature the offline mode. The work on the online mode will be probably too preliminary to be demonstrated.

In such a setting, a key aspect is that of the available knowledge. First, sensor data can be classified based on the kind of knowledge they deliver (e.g. location, movement) and its characteristics (e.g. frequency, accuracy). Then the system has access to geographic data (roads, points of interest), public transportation data (routes and schedules), and finally historical traffic data (traffic jams, schedule delays). These allow creating a *Multimodal Transport Network*. The key notion is that of *admissible multimodal itineraries* that are paths of this network with weights that give the statistics (expected time and variance) of the paths. Finally, the system may rely on the user's personal knowledge such as a history of past itineraries to bias the probabilities of the different routes. Difficulties arise from lack of data (e.g., lack of positioning inside the metro) and from too much data (e.g., combination of possibly conflicting localisation data, overlapping public transportation lines).

For our demonstration, the Multimodal Transport Network is constructed from publicly available transport data of Paris and its neighbourhoods published by different transport agencies and mapping organizations. In terms of transport

BDA 14 octobre 2014, Grenoble-Autrans, France.

network heterogeneity, the region features a multitude of modes for public transport: sub-urban train, metro, bus, tram. These are spread over different transport agencies. The region is also notable for its car-sharing and bike-sharing systems. Smartphone data is recorded via an Android application by multiple users travelling over the Paris region. We demonstrate our system by displaying recognized multimodal itineraries and their estimated likelihood for a single trip of a single-user over an interactive map.

One contribution of our approach is an algorithm that computes a set of most likely itineraries using a combination of particle filtering [9, 10, 28] and a special path sampling algorithm over a Dynamic Bayesian Network [20]. Figure 1 summarizes the different components of our system.

Related work

Activity recognition, that is concerned with determining the actions and goal of one or several agents given a series of observations on their actions and the environment, has gained increased attention over the years in the fields of articial intelligence, robotics, and ubiquitous computing. Our work follows pioneering works published over the last ten years on sensor-based activity recognition, for which accelerometer based approaches are notable [3, 7, 15, 19, 22, 24]. More recently, several researchers have been particularly concerned with *transport mode* identification [14, 18, 25, 29, 30], for which more general approaches can be devised. In these approaches, we are only interested in determining whether the user is currently *stationary, walking, running, cycling* or in a *motor vehicle*. This technology is now available, via their respective APIs, to developers of the two most widespread smartphone platforms [1, 8].

In our setting, the requirement is to determine both transport mode and traveled routes. To serve this purpose, and provide context awareness to travel assistant technologies, location-based activity tracking were developed, that use positional information from an embedded GPS chip to track the user's most frequent locations, travel routines and routes taken [2, 6, 16]. Location-based vehicle tracking has been successfully used to generate accurate schedules and provide information about traffic delays to the rest of the community [4, 26, 27]. On the other hand, our interest is to provide a single user with real-time itinerary-aware applications and an accurate summary of one's routines. To this respect, our work is closely related to the literature in [6, 16], which uses a combination of online activity recognition techniques and location-based map-matching. However, these publications do not distinguish overlapping public transportation routes with different schedules and do not handle long periods of missing GPS observations. Map matching, for which a survey can be found in [23] with more recent results in [17, 21], is concerned with matching a set of positional observations to a path in a given road network. In our case, we cover a more general problem that considers a multimodal transport network. Our multimodal network adds two extra difficulties: timetable management and the fact that two geographic routes may belong to several transport modes and lines.

In itinerary detection, the use of a dynamic Bayesian network is not new. It had to our knowledge never been stressed as far as we do by considering richer transportation data (multimodal distinguishing lines and timetables), with a richer combination of user data: both positional sensors and accelerometer. Another novelty is in the processing of zones with long periods of time with missing GPS observations, which happen frequently in high density urban areas, e.g. with an underground transit system.

Sensor Data

We collected sensor measurements from multiple users equipped with a smartphone and running a logging application. In this section, we quickly overview the different types of sensor measurements recorded in our experiments.

Absolute Positioning. Three different technologies are able to pinpoint the smartphone's position. They come with varying degrees of accuracy, from the most to the least accurate: Satellite Navigation (GPS, GLONASS), Wi-Fi Networks and Cellular Network. These sensors will raise positional events the moment they get a new reading of where the smartphone might be. We use the term *location fix* to refer to such events. Metadata (signal quality, visible satellite count) is recorded as well. We constrain geographic locations to a longitude and latitude pair with no elevation information. Satellite Navigation, besides being the most accurate, is the only one capable of providing speed and bearing as well. Absolute positioning technologies are all characterized by being dependent on wireless transmission of electromagnetic signals. Also, for them to work, they require some kind of static almanac or algorithm that accurately maps signal emitters to a position on Earth or in space. We note that signal unavailability or failure to determine one's position is a very important piece of information. In some cases, we could get a hint of where the user might be, and in many, we will know where the user *cannot be.*

Inertial. Most smartphones have an embedded accelerometer, a gyroscope and a magnetometer sensor. Altogether they provide information regarding the device's movement with respect to an earth's bound frame of reference. Given an initial starting point, orientation and velocity, and sufficient sensor accuracy, the position of the smartphone can be tracked over time. For transport mode detection, we are mostly concerned by accelerometer measurements [14, 18, 25, 29]. However, since the accelerometer measures acceleration within the device's inertial frame, it is measured with respect to the device's orientation. Thus, ideally, one needs to combine information from other sensors to derive a geo-centric orientation. We will not further discuss gyroscope/magnetometer data. Mainly, we will consider that accelerometer data has been augmented, when possible, with gyroscope/magnetometer data. A system that performs this augmentation (sensor-fusion) is found in [13].

Frequency of measurements. Location sensors do not necessarily provide a *location fix* at the requested rate. For example, Cellular/Wi-Fi Network Position sensors do not necessarily raise *fixes* if they deem that the position of the smartphone has not sufficiently changed provided the accuracy of their measurements. Satellite Navigation, which can acquire fixes at 1 Hz irrespective of a position change, can temporarily stop raising *fixes* if the received satellite signals are not strong enough for a sufficient lapse of time.

Accelerometer and other low-level sensors are able to acquire data at an almost fixed-rate, and are not subject to interruptions under a wide array of circumstances. We can

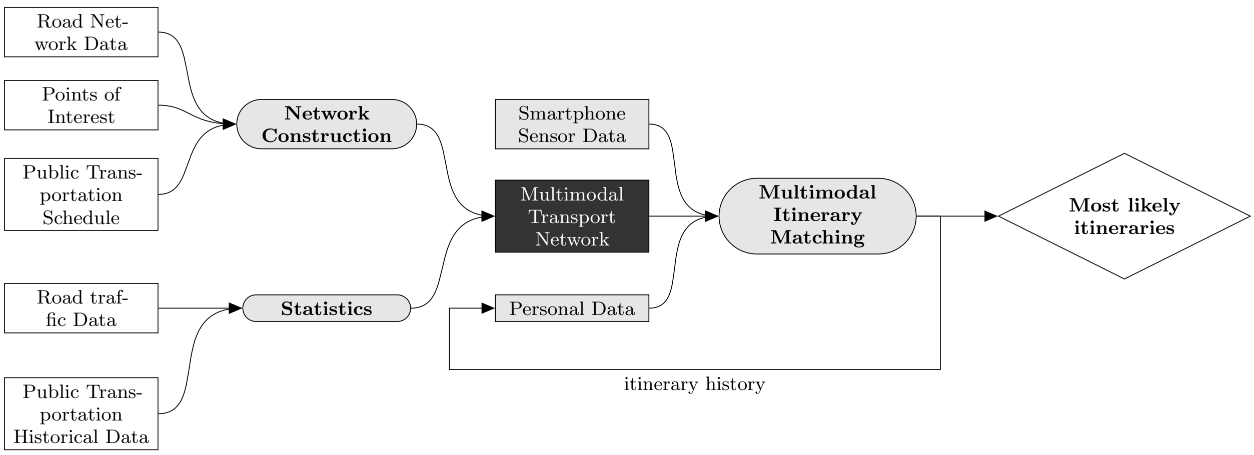

Figure 1: Overall picture: white rectangles are input geographic data, grey ones are user data, dark ones are derived models. Rounded rectangles represent algorithms, and the diamond is the output.

expect such readings at a rate between 50 and 100 Hz. Sensor-fused accelerometer in the geo-centric frame data is thus expected at a rate of at least 50 Hz.

Recording of measurements. Measurements are retrieved from multiple users equipped with an Android smartphone. For privacy considerations, users are asked to press a “Start logging” button at the start each trip, and they are asked to press “Stop logging” at the end of it. During the trip, the application runs in the background, listening to incoming sensor data, and records it in tabular form. Each row represents a single measurement, and is associated a timestamp. The application produces different tabular files for each sensor and type of measurement.

Parisian Multimodal Transport Network

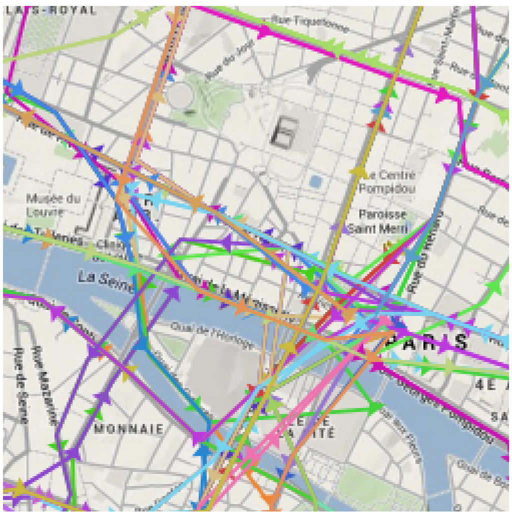

Figure 2: Glimpse of the Parisian public transportation lines, Paris 1er arrondissement

The *Multimodal Transport Network* is constructed with data from various Web sources, including OpenStreetMap (road and infrastructure), RATP & SNCF Open Data (public

Type of data	Count
Roads	$> 10{,}000$
Public transportation lines	375
Bike-sharing stations	1,227
Car-sharing stations	905

Table 1: Multimodal transport network: Paris and neighborhoods

transportation), as well as Velib’ and Autolib’. We distinguish private transport modes (*walking/running, cycling, car, motorcycle*) from public (*bus, metro, tram, train*). For public transport modes, we distinguish different transportation lines (e.g. metro Line 1, Line 2, etc.). Our system is able to construct “admissible multimodal itineraries” that are paths of this network with weights that give the statistics (expected time and variance) of the paths. Similar to [30], mode transitions are only allowed in and out of *human* transport modes (*walking/running, cycling*). Waiting times, e.g. when waiting for a bus at the station, are considered as well. Multimodal Transport Network takes into account the time variability of traffic. As of now, this variability only includes public transportation schedules. More information will be introduced when it is available. For instance, the multimodal transport network will tell (at a particular time of a given day) the average time it would take (and variance) to pick up a bike at the Marguerite de Navarre bike-sharing station, drop it at Grands Boulevards and get on Line 9 of metro.

Time variability of traffic is our first step towards modelling uncertainty of complex events in the Multimodal Transport Network. Link uncertainty, such as missing or modified roads and public transportation lines, although frequent, is not considered in this work. Table 1 shows some key figures of our Multimodal Transport Network. Figure 2 overlays a view of different public transportation lines over a map of Paris at the time of the submission.

Multimodal Itinerary Matching

The system is able to compute the probability that a multimodal itinerary was taken. We model the probability that

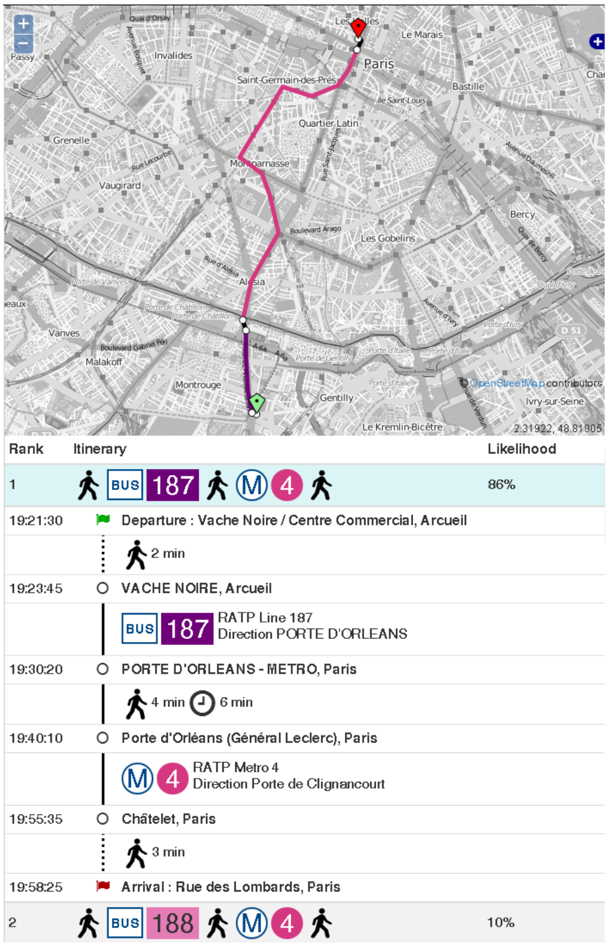

Figure 3: Displaying the most likely itineraries. The itinerary displayed on the map is highlighted in blue in the list.

a *multimodal itinerary* was taken via a Dynamic Bayesian Network [20]. Based on probabilistic observational variables and state transitions, one's current itinerary is represented in the Dynamic Bayesian Network through a state variable that keeps tracks of one's current location in the Multimodal Transport Network. The model integrates a certain number of priors, gathered from the training of local models, verified in practice and extracted from knowledge in previous work [2, 6, 14, 16, 25, 29, 30]. The system performs approximate inference by keeping track of a belief state via particle filtering [5, 9, 12, 16] with a special path sampling algorithm [6, 11].

Demonstration setting

For the demo, user data will be displayed in the form of a list of trips.

For each trip, multiple inferred multimodal itineraries will be displayed in a list, ranked by their likelihood with respect to trip measurements. Selecting an itinerary will overlay it on a map (Fig. 3). When the algorithm does not succeed to infer precise portions of an itinerary, they will be displayed as "grey zones". Raw positional measurements will be overlaid on the map as a baseline. When available, a manually annotated ground truth itinerary may be overlaid for comparison. To demonstrate that the system takes public transportation schedules into account, we will show how an artificially added time offset on raw data produces different results.

Finally, we will display various per-user statistics: frequent itineraries, average departure times and duration of travel. When a user uses multiple itineraries for the same start and endpoints, their statistics may be compared.

References

[1] Android Developers Guide. *Recognizing the User's Current Activity*. 2013. URL: `http://developer.android.com/training/location/activity-recognition.html` (visited on 05/16/2014).

[2] D. Ashbrook and T. Starner. "Using GPS to learn significant locations and predict movement across multiple users". English. In: *Personal and Ubiquitous Computing* 7.5 (2003), pp. 275–286.

[3] L. Bao and S. Intille. "Activity Recognition from User-Annotated Acceleration Data". In: *Pervasive Computing*. Ed. by A. Ferscha and F. Mattern. Vol. 3001. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2004, pp. 1–17.

[4] J. Biagioni, T. Gerlich, T. Merrifield, and J. Eriksson. "EasyTracker: Automatic Transit Tracking, Mapping, and Arrival Time Prediction Using Smartphones". In: *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems*. SenSys '11. Seattle, Washington: ACM, 2011, pp. 68–81.

[5] H. H. Bui, S. Venkatesh, and G. West. "Policy Recognition in the Abstract Hidden Markov Model". In: *J. Artif. Int. Res.* 17.1 (Dec. 2002), pp. 451–499.

[6] J. Chen and M. Bierlaire. "Probabilistic multimodal map-matching with rich smartphone data". In: *Journal of Intelligent Transportation Systems: Technology, Planning and Operations* (2013).

[7] T. Choudhury, S. Consolvo, B. Harrison, J. Hightower, A. Lamarca, L. Legrand, A. Rahimi, A. Rea, G. Bordello, B. Hemingway, P. Klasnja, K. Koscher, J. Landay, J. Lester, D. Wyatt, and D. Haehnel. "The Mobile Sensing Platform: An Embedded Activity Recognition System". In: *Pervasive Computing, IEEE* 7.2 (Apr. 2008), pp. 32–41.

[8] iOS Developer Library. *CMMotionActivity Class Reference*. 2013. URL: `https://developer.apple.com/library/ios/documentation/CoreMotion/Reference/CMMotionActivity_class/Reference/Reference.html` (visited on 05/16/2014).

[9] A. Doucet, N. de Freitas, K. Murphy, and S. Russell. "Rao-blackwellised Particle Filtering for Dynamic Bayesian Networks". In: *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence*. UAI'00. Stanford, California: Morgan Kaufmann Publishers Inc., 2000, pp. 176–183.

[10] A. Doucet, N. Freitas, and N. Gordon. "An Introduction to Sequential Monte Carlo Methods". English. In: *Sequential Monte Carlo Methods in Practice*. Ed. by A. Doucet, N. Freitas, and N. Gordon. Statistics for Engineering and Information Science. Springer New York, 2001, pp. 3–14.

[11] G. Flötteröd and M. Bierlaire. "Metropolis–Hastings sampling of paths". In: *Transportation Research Part B: Methodological* 48 (2013), pp. 53–66.

[12] D. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello. “Bayesian Filtering for Location Estimation”. In: *IEEE Pervasive Computing* 2.3 (2003), pp. 24–33.

[13] E. Foxlin. “Pedestrian tracking with shoe-mounted inertial sensors”. In: *Computer Graphics and Applications, IEEE* 25.6 (Nov. 2005), pp. 38–46.

[14] S. Hemminki, P. Nurmi, and S. Tarkoma. “Accelerometer-based Transportation Mode Detection on Smartphones”. In: *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*. SenSys ’13. Roma, Italy: ACM, 2013, 13:1–13:14.

[15] J. R. Kwapisz, G. M. Weiss, and S. A. Moore. “Activity recognition using cell phone accelerometers”. In: *ACM SigKDD Explorations Newsletter* 12.2 (2011), pp. 74–82.

[16] L. Liao, D. J. Patterson, D. Fox, and H. Kautz. “Learning and inferring transportation routines”. In: *Artificial Intelligence* 171 (2007), pp. 311–331.

[17] Y. Lou, C. Zhang, Y. Zheng, X. Xie, W. Wang, and Y. Huang. “Map-matching for Low-sampling-rate GPS Trajectories”. In: *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. GIS ’09. Seattle, Washington: ACM, 2009, pp. 352–361.

[18] V. Manzoni, D. Maniloff, K. Kloeckl, and C. Ratti. *Transportation mode identification and real-time CO2 emission estimation using smartphones*. Tech. rep. Technical report, Massachusetts Institute of Technology, Cambridge, 2010.

[19] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher. “Activity recognition and monitoring using multiple sensors on different body positions”. In: *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*. IEEE. 2006, 4–pp.

[20] K. P. Murphy. “Dynamic bayesian networks: representation, inference and learning”. PhD thesis. University of California, 2002.

[21] P. Newson and J. Krumm. “Hidden Markov Map Matching Through Noise and Sparseness”. In: *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. GIS ’09. Seattle, Washington: ACM, 2009, pp. 336–343.

[22] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen. “Activity classification using realistic data from wearable sensors”. In: *Information Technology in Biomedicine, IEEE Transactions on* 10.1 (Jan. 2006), pp. 119–128.

[23] M. A. Quddus, W. Y. Ochieng, and R. B. Noland. “Current map-matching algorithms for transport applications: State-of-the art and future research directions”. In: *Transportation Research Part C: Emerging Technologies* 15.5 (2007), pp. 312–328.

[24] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman. “Activity recognition from accelerometer data”. In: *AAAI*. Vol. 5. 2005, pp. 1541–1546.

[25] S. Reddy, M. Mun, J. Burke, D. Estrin, M. Hansen, and M. Srivastava. “Using mobile phones to determine transportation modes”. In: *ACM Transactions on Sensor Networks (TOSN)* 6.2 (2010), p. 13.

[26] A. Thiagarajan, J. Biagioni, T. Gerlich, and J. Eriksson. “Cooperative transit tracking using smart-phones”. In: *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*. SenSys ’10. Zürich, Switzerland: ACM, 2010, pp. 85–98.

[27] A. Thiagarajan, L. Ravindranath, K. LaCurts, S. Madden, H. Balakrishnan, S. Toledo, and J. Eriksson. “VTrack: accurate, energy-aware road traffic delay estimation using mobile phones”. In: *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*. ACM. 2009, pp. 85–98.

[28] S. Thrun, W. Burgard, and D. Fox. *Probabilistic robotics*. MIT press, 2005.

[29] S. Wang, C. Chen, and J. Ma. “Accelerometer Based Transportation Mode Recognition on Mobile Phones”. In: *Wearable Computing Systems (APWCS), 2010 Asia-Pacific Conference on*. Apr. 2010, pp. 44–46.

[30] Y. Zheng, Y. Chen, Q. Li, X. Xie, and W.-Y. Ma. “Understanding transportation modes based on GPS data for web applications”. In: *ACM Transactions on the Web (TWEB)* 4.1 (2010), p. 1.