Uncertainty in Personal Navigation Services

Hassan A. Karimi, Duangduen Roongpiboonsopit,
Piyawan Kasemsuppakorn

(*School of Information Sciences*, *University of Pittsburgh*)
(Email: hkarimi@pitt.edu)

The demand for navigation assistance and advances in several technologies has been paving the way for Personal Navigation Services (PerNavs). As users increasingly rely on PerNavs for navigation assistance, they gain a better understanding of what PerNavs can offer and how they operate. This trend, consequently, will increase the demand for PerNav that can provide high quality solutions. While there have been studies addressing uncertainties associated with selected individual navigation modules, there is a void in the literature addressing the overall uncertainty in PerNavs. In this paper, we discuss uncertainty in PerNavs by analyzing uncertainties associated with each of its modules and how they propagate and impact other modules. A Bayesian network is presented as one possible model to manage (by developers) and communicate (to users) uncertainty in PerNavs.

KEY WORDS

1. personal navigation services.   2. navigation errors.   3. navigation errors management.
4. navigation uncertainty.

1. INTRODUCTION. Navigation systems and services have experienced a tremendous increase in demand over the years. While this can be attributed to the Global Positioning System (GPS) debut in the late 1990s, the continuous trend can be attributed to affordability of GPS receivers, possibility of obtaining high GPS accuracy, availability of map databases in different geographic areas, accessibility to wireless communications, and advances in mobile devices, including cell phones and smartphones. Today, navigation gadgets are common in cars and are widely used by drivers. A decade ago, navigation gadgets were primarily available as *in-car navigation systems* (installed in selected cars by automobile manufacturers) and over time they evolved to *car navigation systems* (portable gadgets which could be used in any car). Modern navigation gadgets are becoming widely available as *navigation services* (provided on cell phones and smartphones, by service providers) and are capable of providing personalized navigation assistance. We call this new trend in navigation Personal Navigation Services (PerNavs). Like other computer systems, PerNavs are susceptible to errors and ambiguities.

While there have been numerous studies on the accuracy of GPS, among other positioning sensors, and map databases and to a certain extent on accuracy of map matching, there is a void in the literature discussing and analyzing the overall

uncertainty in PerNavs. Without a thorough understanding of uncertainty in PerNavs, explaining solutions provided by navigation services to end users is a daunting task and users' confidence dwindles as the intricacies of such services are not appreciated.

In this paper, we first discuss and analyze errors associated with each of the five major modules of a PerNav: map database, geocoding, positioning, map matching, and routing and direction. While a map database provides the operational foundation for the other four modules in a PerNav, its errors are propagated to other modules and each of the other four modules contributes some errors to the overall navigation services. We further provide a tool based on a Bayesian network (BN) model to help developers and users in understanding uncertainty in PerNavs.

The contributions of this paper are: (a) identification of sources of errors in PerNavs, (b) analysis of uncertainty in PerNavs, and (c) a model to manage uncertainty in PerNavs. The structure of the paper is as follows. Sections 2–6 discuss uncertainty sources in map database, geocoding, positioning, map matching, and routing/direction. Section 7 presents a BN model to manage uncertainty in PerNavs. The paper ends with a summary in Section 8.

2. MAP DATABASE. As a core module of PerNavs, a road network database supports almost every other module of a navigation system (Skog and Handel, 2009). The map matching module relies on road network data to determine a best estimate of current vehicle's position. The geocoding module uses road network data to interpolate the coordinates of a given address. The routing/direction module relies on road network data to calculate an optimal route and generate directions on it.

Road network data used in PerNavs includes geometries, topologies, and attributes, represented in the vector data model. A road network database containing basic attributes (e.g., street name, address range, length, street type) for geocoding and routing purposes is called a *routable map*. A road network database containing a complete set of attributes (e.g., orientation, landmarks, real-time traffic) for all navigation functions is called a *navigable map*. Today, many digital map data providers create and provide road network data and other map related content. Generally, map data providers can be grouped into three categories: (a) government and non-profit organizations, (b) commercial mapping companies, and (c) community mapping projects. Map data providers and their products, navigable and routable maps, are shown in Figure 1.

2.1. *Sources of Errors in Map Databases.* Road network data errors come from multiple sources during the map generation process which has the following steps: data acquisition, data automation, and map creation, as summarized in Figure 2.

Data acquisition is an extensive task and error prone. In addition to common acquisition techniques for road network data (i.e., field survey, paper map conversion, and satellite image feature extraction), a newer approach for collecting road data in a large geographical extent for creating network databases is through mobile mapping systems (Karimi and Grejner-Brzezinska, 2004).

Paper maps can either be digitized or scanned. Typical accuracy of paper map digitization ranges between 1·5 m to 15 m from its true position (Fekpe et al., 2003). Typical errors in scanning come from the map itself and the tracing method (Chang,
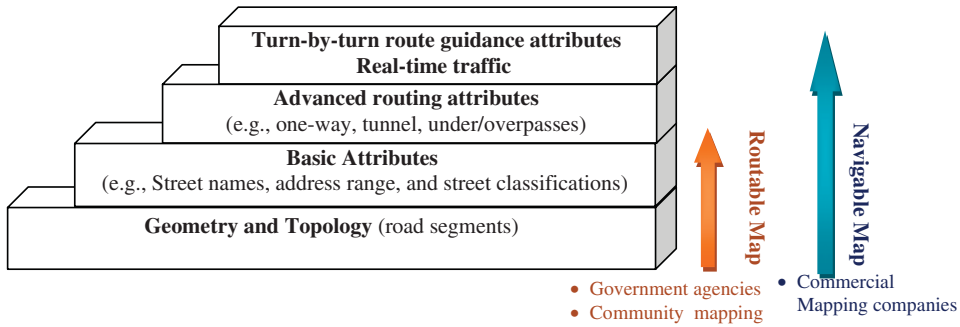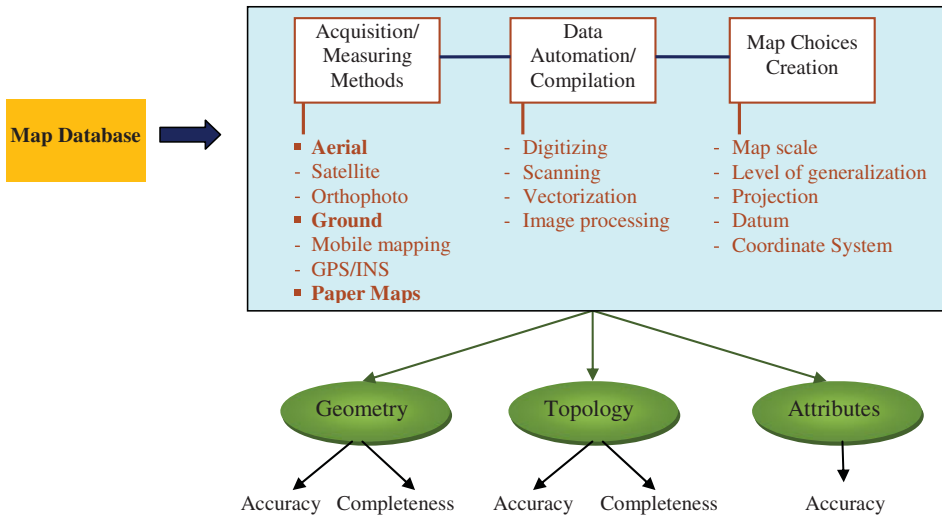
Figure 1. Road network databases.

Figure 2. Sources of errors in digital maps.

2010). High-resolution aerial and satellite images are alternative sources of data for extracting road networks and updating road network data. Typical accuracy of aerial photographs, taken by high quality stereo plotters, ranges between 0·08 m and 0·13 m with image scale of 0·025 : 30·48 m (Fekpe et al., 2003).

The map creation process involves a series of choices on how real-world features should be represented in a digital map. Each decision may introduce uncertainties to the overall process. The steps in the map creation process include choice of map scale, level of generalization, projection, datum, and coordinate system. Maps suitable for car navigation systems and Intelligent Transportation System (ITS) applications are large-scale, about 1:5,000 for urban areas and 1:10,000 for rural areas (Zhao, 1997). A choice of generalization usually distorts details of map features; for example, the street centrelines model might eliminate small curves of actual road shapes. The choice of map projection, horizontal and vertical datum, and coordinate system may introduce some distortions and distance measurement errors.

2.2. *Measuring Map Database Errors*. Road network data quality can be measured using various metrics, which can be classified into three groups: geometry

(accuracy, completeness), topology (accuracy, completeness), and attribute (accuracy). Geometrical accuracy refers to the closeness of road coordinates in the database to true positions of the actual roads. A road network database is considered to be geometrically complete if it contains all the actual road network features within the underlying geographic extent. Topological accuracy refers to the degree to which geographic features represent correct connections. A road network database is considered to be topologically complete if it contains all the actual intersections within the underlying geographic extent. Attribute accuracy refers to the discrepancy between attribute data in the database and true values of the attributes.

There are no standards available for measuring geometrical completeness, topological accuracy, and topological completeness parameters. However, these three parameters play a major role in road network quality and uncertainty in PerNavs. One possible method of measuring these parameters is to use an external evaluation methodology based on image processing techniques, such as the metrics proposed by Wiedemann (2003).

3. GEOCODING. Geocoding is the process of assigning geographic coordinates to a given place name by comparing its description to the descriptions of location-specific elements in the reference database. For a comprehensive review of the geocoding process refer to Goldberg et al. (2007), Rushton et al. (2006), and Zandbergen (2008, 2009). Geocoding in PerNavs is required to locate points of interest (POIs) and origin-destination pairs for route planning.

PerNavs allow users to specify a desired destination in three ways: placing a point on the map, entering an address/intersection name, or selecting a POI from lists of directories. Usually, POIs in PersNavs are either pre-geocoded or geocoded on the fly. In the pre-geocoding method, coordinates of POIs or major landmarks (e.g., airport, parks, parking lots, and shopping centres) are stored in a database. On-the-fly methods find coordinates of an address at the time the address is provided.

3.1. *Sources of Errors in Geocoding.* Geocoding uncertainties are mainly associated with choices of techniques (street geocoding, rooftop geocoding), algorithms and reference databases. (Roongpiboonsopit and Karimi, 2010a, 2010b). In this paper, uncertainties associated with the street geocoding technique, since it is widely used in PerNavs, are focused. Figure 3 summarizes sources of errors for street geocoding.

3.1.1. *Reference Database.* Street geocoding uses road network data in a reference database. Following the US address model (i.e., house number, street prefix, street name, street suffix, postal codes, city, state, and country), the attributes that most likely introduce errors to geocoding results are street names, house number ranges, and polarity (i.e., even/odd house numbers assigned to left/right side of street) since any changes on these may not be reflected in reference databases immediately. A reference database containing many of problematic attributes can cause low match rates and possibly low positional accuracy (Karimi et al, 2004).

Geometrical accuracy of street centrelines and intersection nodes directly influences positional accuracy of geocoded results. Street resolution, which is the sampling frequency of shape points along each segment, is another source of errors. A low resolution segment results in a coarse representation of the actual street and a rough estimate of the total length of the segment.
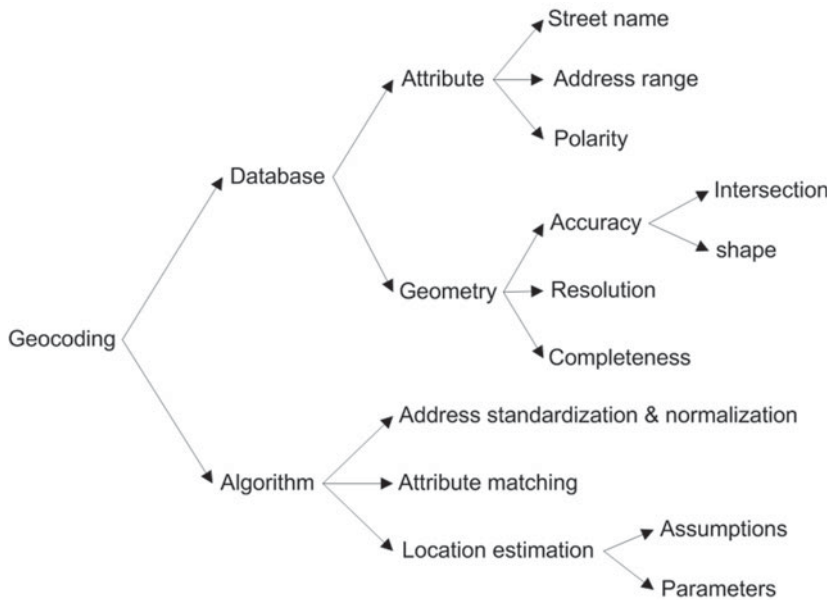
Figure 3. Sources of errors in street geocoding.

Completeness of street centrelines in reference databases is another important factor impacting match rates. Even when the best geocoding algorithm is employed, missing major street centrelines associated with many addresses in the reference database can cause low match rates.

3.1.2. *Geocoding Algorithm.* A geocoding algorithm involves three steps: address standardization and normalization, attribute matching, and location estimation. Approaches, assumptions, and parameters in each step introduce uncertainties in geocoding results. Karimi et al. (2004) showed that using the same reference database in different geocoding algorithms may not have a significant impact on results.

Address standardization and normalization converts an input address to a format that is understandable and compatible with the reference database. The geocoder must be robust and have provisions for unanticipated errors associated with input addresses. A simple standardization and normalization method (e.g., simple token parsing) leads to lower attribute matches than advanced methods, such as hidden Markov models, that can cope with many types of misspellings and misplacements rather well (Christen et al., 2004, Yang et al., 2004).

Attribute matching is the first step in geocoding that finds the correct street for interpolation. Different approaches, such as word stemming, Soundex, and relaxation of matching requirements, can be applied to improve finding a match in the reference database (Churches et al., 2002, Drummond, 1995, and Nicoara, 2005). A scoring scheme for quantifying quality of matches is applied which usually returns a score ranging between 0 to 100. A minimum matching score for examining street centreline candidates can be assigned. Assigning an appropriate minimum score is important because a too low score can result in higher match rates but lower positional accuracy of the output (Bonner et al., 2003, Rushton et al., 2006).

Location estimation aims to interpolate the coordinates of the parsed address using the geometry and attributes of the selected street segment. In addition to the values of attributes (i.e., address range and polarity) necessary for interpolation (Cayo and Talbot, 2003, Levine and Kim, 1998), the algorithm has a set of underlying assumptions and internal parameters defined to improve the positional accuracy of the output (Zandbergen, 2009).

Two common assumptions in street geocoding are parcel homogeneity and address range existence. The parameters of an interpolation algorithm that can impact positional accuracy are side offsets and end offsets. An interpolation algorithm assumes that all addresses are located at a certain distance away from a street centerline and end nodes. Applying unreasonable offsets can magnify distance errors. However, Zandbergen (2007) reported minor differences on the positional error distribution when offset values were varied between 0 to 50 m.

3.2. *Measuring Geocoding Errors.* Geocoding quality is commonly measured using two metrics: match rates and positional accuracy. Match rate is the percentage of input addresses that can be geocoded. Positional accuracy is a statistical measure that provides the degree of conformance between the geocoded points and the actual location of the addresses.

4. POSITIONING. The positioning module of a PerNav is responsible for continually estimating geographic coordinates of a vehicle's current location. Coordinates obtained by the positioning module are used to locate the vehicle's position on a map by the use of map matching techniques. The map-matched points then allow follow-up functionalities in PerNavs, including computing a route from the vehicle's current location to a destination, re-calculating a route, giving turn-by-turn directions, and searching nearby POIs.

Currently, several positioning technologies are available for location-based services and applications, including Global Navigation Satellite System (GNSS), Dead Reckoning (DR), Wi-Fi, cellular networks, Ultra Wide Band (UWB), Radio Frequency Identification (RFID), and Bluetooth positioning (Retscher and Kealy, 2006, Rizos, 2005). Of available positioning technologies, GNSS is predominantly utilized in PerNavs. However, GNSS suffers from losses of satellite signals in difficult areas such as indoors, urban canyons, heavy canopies, and tunnels. To overcome GNSS drawbacks and increase continuity and availability of positioning solutions, GNSS is often augmented by DR systems in PerNavs (Hide et al., 2006, Kao, 1991, and Zhao et al., 2003). In this paper we focus on the uncertainties associated with GNSS as it is commonly employed in PerNavs.

GNSS is an absolute radio-navigation positioning systems and currently the only fully operational GNSS is GPS. Russia's GLONASS, Europe's Galileo, and China's Compass will be available with full operation in the near future. For more comprehensive details on GNSS refer to Drane and Rizos (1998), Hofmann-Wellenhof et al. (2008), and Kaplan and Hegarty (2006).

4.1. *Sources of Errors in Positioning.* Figure 4 summarizes major sources of errors in GNSS positioning. GNSS accuracy depends on the level of measurement errors and satellite-receiver geometry. Some of these measurement errors could be reduced or eliminated by GNSS augmentation techniques (e.g., differential GNSS). Poor satellite geometry, on the other hand, cannot be corrected.
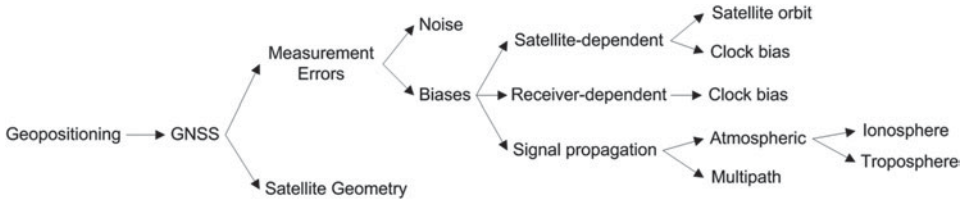
Figure 4. Sources of errors in positioning.

Measurement errors can be categorized into three groups: satellite-dependent, receiver-dependent, and signal propagation. Satellite-dependent biases consist of orbital bias and satellite clock bias. Receiver-dependent bias is primarily clock error due to the use of inexpensive quartz crystal oscillators, which can cause the difference between the receiver clock time and the satellite clock time. Signal propagation biases occur due to atmospheric and multipath effects. The ionosphere and troposphere cause slower velocity for the propagating signal than when it travels in outer space.

The multipath effect occurs whenever the signal takes an indirect path, reflecting from nearby objects, rather than the direct path from satellite to receiver. The delay can cause error for code pseudorange between 10 to 20 m and may increase to 100 m in severe areas. The multipath effect in PerNavs is more challenging due to varying surroundings from place to place, and introduces the most significant bias to the final position solution.

To combine contributions by different sources of measurement errors described earlier, user equivalent range error (UERE) is defined. Assuming that individual errors are uncorrelated, UERE associated with a satellite is computed as a square root of the summed squares of the individual errors (σ_{UERE}). For further reading on the amount of each measurement error and UERE for different GNSS types refer to Hofmann-Wellenhof et al. (2008) and Kaplan and Hegarty (2006).

Satellite geometry is an amplification factor of the measurement errors that cannot be corrected by any means. The parameter for determining satellite geometry is dilution of precision (DOP) (Langley, 1999). Based on different components used in DOP calculation, several definitions of DOP are possible. The most common DOPs for ground applications are positional DOP (3D coordinates) and horizontal DOP (2D coordinates).

4.2. *Measuring Positioning Errors.* To determine positioning performances of GNSS, four metrics, termed Required Navigation Performance (RNP), are defined: accuracy, availability, continuity, and integrity (Ochieng et al., 1999). Accuracy is a statistical measure that provides a degree of conformance between estimated position and true location of an object. Availability refers to percentage of time that the positioning information is available to use. Continuity is the capability to generate a stream of positions without outage during an intended period of operation. Integrity refers to the level of trust in the correctness of position information.

5. MAP MATCHING. Map matching is the process of determining current location of the vehicle on the road network after obtaining its geographic coordinates via the positioning module. In general, map matching is a two-step process. In the first step a correct road segment that the user is travelling on is
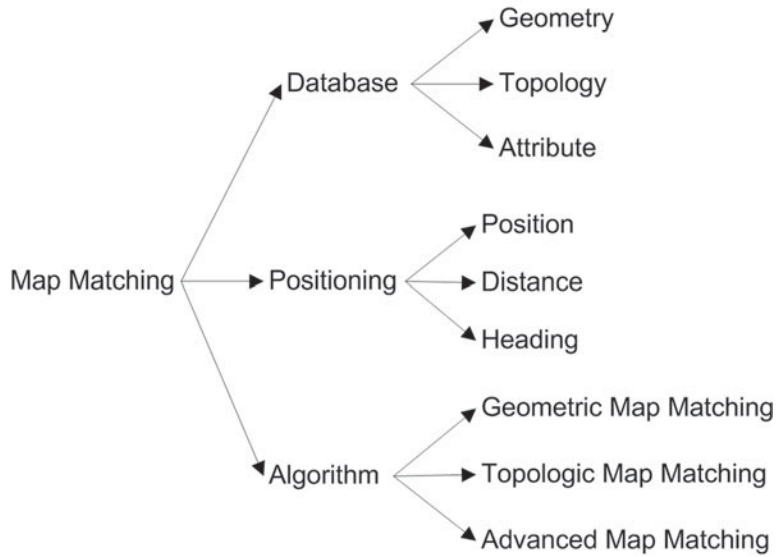
Figure 5. Sources of errors in map matching.

identified and in the second step a location on the identified segment is pinpointed. A map matching algorithm generally utilizes information available onboard, such as position data and road network. This information, as discussed earlier, is imperfect and contains errors.

The level of map-matched results depends on how the algorithm is designed to tolerate uncertainties of the information sources it uses (i.e., positioning and road network). Advanced map matching algorithms can usually handle most ambiguous cases by reducing errors of map-matched points. Karimi et al., (2006) provided a methodology to predict and evaluate performance of different map matching algorithms.

5.1. *Sources of Errors in Map Matching.* Map matching uncertainties come from three main sources: positioning module, digital road network, and map matching algorithms, as summarized in Figure 5. In urban areas where road networks are dense, most map matching algorithms cannot correctly identify correct road segments due to high uncertainties of GNSS positions. For instance, a field test in London found that the GPS error was more than 50 m (Zhao et al., 2003), and in Hong Kong it was more than 80 m (Chen et al., 2003). Even though DR may be used to augment GNSS in PerNavs, drift errors may arise if the system is not calibrated for a period of time. Using the drifted DR points can significantly impact map-matched results. Distance travelled between two points is used by some map matching algorithms to estimate points on the selected segment. Thus, high uncertainty of distance information, either obtained from odometer sensors or derived from GNSS points, may cause errors in map-matched positions. Heading information can be used to identify a road segment, select the next travelling road segment only at the intersection, or determine new candidate segments whenever new points are obtained. Variation in heading may impact map matching solutions depending on how the algorithm utilizes this information. Chen et al. (2005) showed that GPS position

errors contribute 4% of mismatching positions, 22% of which are caused by distance measurement errors and 50% caused by heading errors.

5.2. *Measuring Map-Matching Errors*.   Quality of map-matching results can be determined by two metrics: identification rate (IR) and map-matched accuracy. IR is the ratio between the number of correctly identified segments to the total number of segments in the underlying trajectory. Map-matched accuracy is a statistical measure that provides a degree of conformance between the map-matched point and the true location of the vehicle.

6. R O U T I N G / D I R E C T I O N.   Routing and direction in PerNavs compute user-preferred routes and step-by-step instructions on how to travel on routes, respectively. Typical route preferences are shortest distance, fastest time, avoid highway, and avoid toll (Pang, 2002). Routing is an optimization problem that searches a road network for best routes between pairs of addresses. Several algorithms for solving routing problem exist including Dijkstra's algorithm (1959) and Floyd-Warshall's algorithm (Pemmaraju and Skiena 2003). Generally, routing algorithms are either exact or heuristic. Exact routing algorithms can guarantee best solutions (routes) but with slow time complexity. Heuristic algorithms are an alternative to exact algorithms typically reduce the solution space in order to improve time complexity; for example, A* (Hart et al., 1972) and ORCA (Karimi, 1996). One drawback with heuristic algorithms is that they do not guarantee best solutions (Johnsonbaugh and Schaefer, 2003).

Research in direction generation has been studied mainly as a field in cognitive science, especially in wayfinding. An example of direction generation is GUARD (Generation of Unambiguous, Adapted Route Directions) by Richter (2007) that automatically determines references to different types of landmarks into context specific route instructions. Generally, three types of information are used in PerNavs for generating directions: street names, street length, and orientation at decision points.

6.1. *Sources of Errors in Routing/Direction*.   Uncertainty in routing refers to differences between routes computed and those requested by users. Uncertainty in direction refers to directions that contain incorrect, inaccurate, or outdated information. The primary sources of errors in routing/direction are summarized in Figure 6.

6.1.1. *Routing*.   In routing, weight function computes weights (costs), based on user's preferences, for road segments. Considering that weight functions utilize attributes in map databases, directly or indirectly, their results depend on map database errors. For example, with travel time as a preferred criterion, erroneous road segment speed limits will cause errors in calculating weights.

Route computation consists of three main steps: obtaining origin position, obtaining destination position, and computing route between origin and destination. The vehicle's position in PerNavs is usually considered as the origin, which is obtained from the map matching module and thus susceptible to errors of map matching. Destination can be obtained through the geocoding module, and is thus susceptible to errors of geocoding and map database.

Heuristic routing algorithms may introduce uncertainties which are not issues in exact the routing algorithm. In addition, geometrical, topological, and attribute
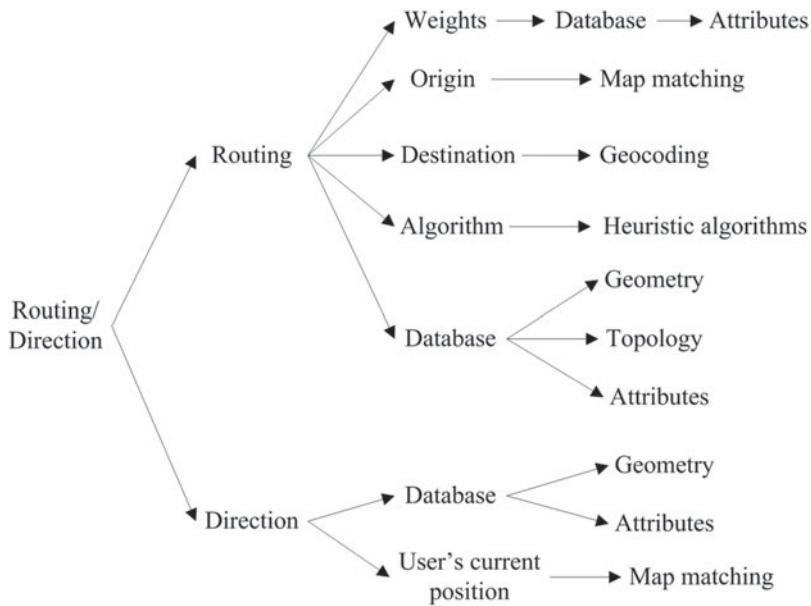
Figure 6. Sources of errors in routing/direction.

errors in road network databases can cause errors in routing solutions; for example, missing road segments, incorrect orientation (one-way or two-way) of road segments, or missing intersection nodes.

6.1.2. *Direction.* Direction generation in PerNavs involves two main steps: generating direction from origin to destination and retrieving turn-by-turn instruction in real time based on the vehicle's position. Uncertainties of direction mainly associate with road attributes. The main sources of errors in the first step of direction generation are segment name and segment length. The sources of errors in the second step of direction generation are associated with the vehicle's position determination at each time epoch using the map matching module and distance estimation to the next decision point based on the vehicle's position, speed, and current road segment length.

6.2. *Measuring Routing/Direction Errors.* To measure errors of the routing and direction modules, we define two sets of parameters, one for measuring route errors and one for direction errors.

6.2.1. *Route Error Measures.* Three parameters for route uncertainty are: weight error (w_e), route precision (R_{prec}) and route completeness (R_{comp}), which can be measured from individual computed routes.

Weight error is the average of normalized differences between weights of segments in the computed route and weights of segments in the true route. True routes are the baseline routes preferred by the users. Weight error can be calculated as follows:

$$e_w = \frac{1}{n} \sum_{i=1}^{n} \frac{w_{ci} - w_{ti}}{w_{ti}} \tag{1}$$

where w_{ci} is the weight on the ith segment of the computed route, w_{ti} is the weight on the ith segment of the true route, and n is the number of segments in the true route.

The segments of the computed routes that are not contained in the true route are discarded since they cannot be compared against the baseline.

Route precision is a measure of exactness of the computed route comparing to the optimal route (the true route). It is defined as the ratio between the number of common road segments in the computed route and in the actual optimal route and number of road segments in the actual optimal route, calculated as follows:

$$R_{prec} = \frac{|(P_c \cap P_o)|}{|P_o|} \tag{2}$$

where P_c is the set of road segments in the computed route, P_o is the set of road segments in the optimal route.

Route completeness is a measure of completeness of the computed route compared to the optimal route. It is defined as the ratio between the number of common road segments in the computed route and in the actual optimal route and number of road segments in the computed route, calculated as follows:

$$R_{comp} = \frac{|(P_c \cap P_o)|}{|P_c|} \tag{3}$$

Thus, if $R_{prec} = 1$ and $R_{comp} = 1$, the computed route is the same as the optimal route; otherwise, there is uncertainty associated with the computed route.

6.2.2. *Direction Error Measures.* The parameters for direction uncertainty are: distance error (e_d), topological error (e_t), and attribute (landmark) error. In order to compute these parameters, the computed route is assumed to be correct and the true direction along the given computed route is used as the baseline direction.

Distance error is the total sum of differences between distances provided in the computed direction and actual distances among entities (e.g., between user's location and landmarks). It is calculated as follows:

$$e_d = \frac{1}{n} \sum_{j=1}^{n} \frac{(d_{cj} - d_{tj})}{d_{tj}} \tag{4}$$

where d_{cj} is the jth distance provided in the direction, d_{tj} is the actual j^{th} distance in the computed route, and n is the number of times the distances are provided in the direction.

Topological error is the difference between the number of turns provided in the direction and the number of true turns in the computed route. It is calculated as follows:

$$e_t = \frac{T_c}{T_t} \tag{5}$$

where T_c is the total of number of turns provided in the direction and T_t is the total number of true turns on the computed route.

Attribute (landmark) error is the discrepancy between the landmarks provided in the direction and the actual landmarks on the computed route. It is calculated as follows:

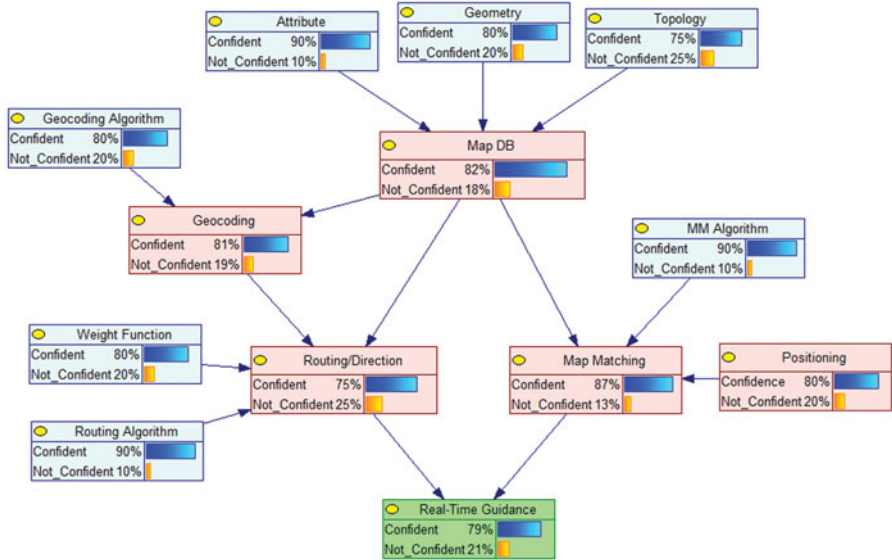$$L_c = \{l_{c1}, l_{c2}, \ldots, l_{cm}\} \tag{6}$$

Figure 7. A Bayesian network for errors in PerNavs with initial probabilistic values.

$$L_t = \{l_{t1}, l_{t2}, \ldots, l_{tn}\} \tag{7}$$

$$L_c \subseteq L_t \text{ and } L_c \cap L_t = L_c \tag{8}$$

where L_c is the set of computed landmarks and L_t is the set of actual landmarks. All elements in L_c must exist in L_t, otherwise the information about the landmarks in the direction is incorrect.

7. NAVIGATION UNCERTAINTY MODEL. This section presents an understanding of uncertainty in navigation by analyzing the impact of uncertainty in each module on overall uncertainty in PerNavs. Of the possible approaches for analyzing uncertainty in navigation, a Bayesian Network (BN) which is a probabilistic graphical model, is employed to analyze uncertainty propagation in a PerNav. The objective of employing a model like BN is to provide PerNav developers with a tool for explaining the performances of their products to end users, assisting them in making appropriate navigation decisions. The BN structure for PerNavs, called Navigation BN (Nav-BN), is shown in Figure 7. Nav-BN is composed of thirteen variables, represented by nodes, which are classified into three groups: specific, modular, and overall. Specific variables represent uncertainties associated with a particular navigation module. Modular variables represent the overall uncertainty of each navigation module. Overall variable represents the total uncertainty in a PerNav. In this particular design, seven nodes are specific variables, five nodes are modular variables, and one node is overall variable. To simplify the network, variables are provided with two possible states: *confidence* and *not_confidence*. Each node is defined as follows:

- *Attribute* is the level of confidence in attribute data, stored in the map database, that meets the required accuracy;

Attribute	Confident				Not_Confident			
Geometry	Confident		Not_Confident		Confident		Not_Confident	
Topology	Confident	Not_Confid...	Confident	Not_Confid...	Confident	Not_Confid...	Confident	Not_Confid...
Confident	0.95	0.8	0.6	0.3	0.85	0.65	0.4	0.2
Not_Confident	0.05	0.2	0.4	0.7	0.15	0.35	0.6	0.8

Figure 8. An example of a conditional probability table (CPT) for the Map DB variable.

- *Geometry* is the level of confidence in geometrical data, stored in the map database, that meets the required accuracy and completeness;
- *Topology* is the level of confidence in topological data, stored in the map database, that meets the required correctness and completeness;
- *Map DB* is the overall confidence in map database influenced by *Geometry*, *Topology*, and *Attributes*;
- *Geocoding Algorithm* is the level of confidence in geocoding algorithm that generates correct results;
- *Geocoding* is the overall confidence in the geocoding module influenced by *Map DB* and *Geocoding Algorithm*;
- *Positioning* is the overall confidence in the positioning module;
- *Map Matching* is the overall confidence in map matching, influenced by *Map DB*, *Positioning*, and *MM Algorithm*;
- *Weight Function* is the level of confidence in the weight function that produces desired costs meeting users' preferences;
- *Routing Algorithm* is the level of confidence in routing algorithm that produces routes satisfying users' preferences;
- *Routing/Direction* is the overall confidence in computed route and direction, influenced by *Map DB, Weight Function, Routing Algorithm*, and *Geocoding*;
- *Real-Time Guidance* is the overall confidence in real-time navigation performance that a PerNav provides correct and satisfactory solutions.

Nav-BN requires prior information tables (PIT), or knowledge, in the form of probabilities to indicate possibilities of the confidence levels for the specific variables and the Positioning variable. For example, a PerNav employs a positioning sensor that can estimate acceptable positions with 80% confidence, as shown in Figure 7 at the Positioning node. Nav-BN also requires conditional probability tables (CPTs) for the modular and overall variables for representing changes on the confidence levels given causes. Figure 8 shows an example of CPT for Map DB variable with initial values. Generally, there are two possible approaches to estimate values in PITs and CPTs, which are estimated from collected data and expert judgment. Collected data used for measuring performances of each module during the design and evaluation process of PerNavs can be used to estimate the probabilistic values of both PITs and CPTs. Alternatively, probabilistic values can be estimated by learning from users' experiences using PerNavs. Figure 7 shows estimated probabilistic values of a Nav-BN for a PerNav based on the authors' judgment. The performance confidence of this particular PerNav is, on average, 79%, which is directly influenced by the uncertainty associated with the map matching and the route/direction modules.

Nav-BN can benefit both PerNav developers and end users in several ways. Nav-BN can help developers better understand uncertainty associated with each module and the overall navigation performance at the Real-Time Guidance node. They can also infer a confidence level in each module if the confidence of Real-Time Guidance

is fixed as a requirement. The inferred confidence level of each module can assist developers in choosing suitable database, positioning sensors and algorithms. Nav-BN benefits end users by helping them make appropriate navigation decisions. Nav-BN can provide periodic estimates of navigation performance by observing and measuring the confidence level of position data in real time. For example, if the level of confidence in positioning solutions is known through errors of estimated co-ordinates (x, y), the map matching module can immediately update the confidence results accordingly, which in turn influences the overall real-time guidance confidence.

8. SUMMARY. In this paper we discussed the importance of understanding uncertainty in navigation by developers and end users. We discussed and analyzed uncertainties associated with each of the five modules (map database, geocoding, positioning, map matching, routing/direction) of PerNavs and how uncertainties of each would propagate to some or all other modules. Uncertainties of each module were analyzed and their sources were discussed. Clearly, a map database is the foundation of a PerNav and its uncertainties propagate to all other modules which impacts subsequent processes. We also presented a Bayesian network model, called Nav-BN, as one possible approach for capturing each module's uncertainty and realizing the impact of uncertainty propagation through modules.

Even though this paper has addressed several issues of uncertainties in navigation, many open issues still remain for future research. One area of future research is to experiment with the presented Bayesian network model to find its suitability and compare it with other possible approaches. While this research area will be of interest to PerNav developers, exploring means of communicating uncertainty in navigation to PerNav end users is another research topic.

REFERENCES

Bonner, M. R., Han, D., Nie, J., Rogerson, P., Vena, J. E. and Freudenheim, J. L. (2003). Positional accuracy of geocoded addresses in epidemiologic research. *Epidemiology*, **14**, 408–412.

Cayo, M. R. and Talbot, T. O. (2003). Positional error in automated geocoding of residential addresses. *International Journal of Health Geographics*, **2**, 1–12.

Chang, K.-T. (2010). *Introduction to geographic information systems* (5th ed.). Mc-Graw Hill.

Chen, W., Li, Z., Yu, M. and Chen, Y. (2003). Integrated vehicle navigation system for urban applications. *Preceedings of the 7th International Conference on Global Satellite Navigation Systems* (*GNSS*), Graz, Austria.

Chen, W., Li, Z., Yu, M. and Chen, Y. (2005). Effects of sensor errors on the performance of map matching. *Journal of Navigation*, **58**, 273–282.

Christen, P., Churches, T. and Willmore, A. (2004). A probabilistic geocoding system based on a national address file. *Proceedings of the 3rd Australasian Data Mining Conference*, Cairns, Australia.

Churches, T., Christen, P., Lim, K. and Zhu, J. X. (2002). Preparation of name and address data for record linkage using hidden Markov models. *BMC Medical Informatics and Decision Making*, **2**, 1–16.

Dijkstra, E. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, **1**, 269–271.

Drane, C. and Rizos, C. (1998). *Positioning systems in intelligent transportation systems*. Artech House Inc.

Drummond, W. J. (1995). Address matching: GIS technology for mapping human activity patterns. *Journal of the American Planning Association*, **61**, 240–251.

Fekpe, E. S., Windholz, T., Beard, K. and Novak, K. (2003). Quality and accuracy of positional data in transportation (No. 506): NCHRP.

Goldberg, D. W., Wilson, J. P. and Knoblock, C. A. (2007). From text to geographic coordinates: the current state of geocoding. *Journal of the Urban and Regional Information Systems Association*, **19**, 33–46.

Hart, P., Nilsson, N. and Raphael, B. (1972). Correction to a formal basis for the heuristic determination of minimum cost paths. *ACM SIGART Bulletin*, 28–29.

Hide, C., Moore, T., Hill, C. and Park, D. (2006). Low cost, high accuracy positioning in urban environments. *The Journal of navigation*, **59**, 365–379.

Hofmann-Wellenhof, B., Lichtenegger, H. and Wasle, E. (2008). *GNSS-global navigation satellite systems: GPS, GLONASS, Galileo, and more*. Springer-Verlag Wien.

Johnsonbaugh, R. and Schaefer, M. (2003). *Algorithms*. Prentice-Hall, Inc.

Kao, W. (1991). Integration of GPS and dead-reckoning navigation Systems. *Preceedings of IEEE Vehicle Navigation and Information Systems Conference* (*VNIS '91*), 635–643.

Kaplan, E. D. and Hegarty, C. J. (2006). *Understanding GPS: principles and applications* (2nd ed.). Artech house.

Karimi, H. and Grejner-Brzezinska, D. (2004). GQMAP: Improving Performance and Productivity of Mobile Mapping Systems through GPS Quality of Service. *Cartography and Geographic Information Science*, **31**, 167–177.

Karimi, H. A. (1996). Real-time optimal-route computation: A heuristic approach. *Intelligent Transportation Systems Journal*, **3**, 111–127.

Karimi, H. A., Conahan, T. and Roongpiboonsopit, D. (2006). A methodology for predicting performances of map-matching algorithms. *Preceedings of the 6th International Symposium on Web and Wireless Geographical Information Systems* (*W2GIS 2006*), Hong Kong, 202–213.

Karimi, H. A., Durcik, M. and Rasdorf, W. (2004). Evaluation of uncertainties associated with geocoding techniques. *Computer-Aided Civil and Infrastructure Engineering*, **19**, 170–185.

Langley, R. B. (1999). Dilution of precision. *GPS World*, **10**, 52–59.

Levine, N. and Kim, K. E. (1998). The location of motor vehicle crashes in Honolulu: a methodology for geocoding intersections. *Computers, Environment and Urban Systems*, **22**, 557–576.

Nicoara, G. (2005). Exploring the geocoding process: a municipal case study using crime data.

Ochieng, W. Y., Shardlow, P. J. and Johnston, G. (1999). Advanced Transport Telematics Positioning Requirements: An Assessment of GPS Performance in Greater London. *Journal of Navigation*, **52**, 342–355.

Pang, G. K. H., Takashi, K., Yokotab, T., and Takenaga, H. (2002). Intelligent route selection for in-vehicle navigation systems. *Transportation Planning and Technology*, **25**, 175–213.

Pemmaraju, S. and Skiena, S. (2003). *Computational discrete mathematics: combinatorics and graph theory with Mathematica*. Cambridge University Press.

Retscher, G. and Kealy, A. (2006). Ubiquitous Positioning Technologies for Modern Intelligent Navigation Systems. *Journal of Navigation*, **59**, 91–103.

Richter, K. (2007). A uniform handling of different landmark types in route directions. *Spatial information theory*, **4736**, 373–389.

Rizos, C. (2005). Trends in geopositioning for LBS, navigation and mapping. *Preceedings of International Symposium and Exhibition on Geoinformation*, Penang, Malaysia,

Roongpiboonsopit, D. and Karimi, H. A. (2010a). Comparative evaluation and analysis of online geocoding services. *International Journal of Geographical Information Science*, **24**, 1081–1100.

Roongpiboonsopit, D. and Karimi, H. A. (2010b). Quality assessment of online street and rooftop geocoding services. *Cartography and Geographic Information Science*, **37**, 301–318.

Rushton, G., Armstrong, M., Gittler, J., Greene, B., Pavlik, C., West, M., et al. (2006). Geocoding in Cancer Research A Review. *American journal of preventive medicine*, **30**, S16–S24.

Skog, I. and Handel, P. (2009). In-car positioning and navigation technologies – a survey. *IEEE Transactions on Intelligent Transportation Systems*, **10**, 4–21.

Wiedemann, C. (2003). External evaluation of road networks. *ISPRS Archieves*, **XXXIV**, 93–98.

Yang, D., Bilaver, L. M., Hayes, O. and Goerge, R. (2004). Improving geocoding practice : evaluation of geocidng tools. *Journal of Medical Systems*, **28**, 361–370.

Zandbergen, P. A. (2007). Influence of geocoding quality on environmental exposure assessment of children living near high traffic roads. *BMC Public Health*, **7**, 37–50.

Zandbergen, P. A. (2008). A comparison of address point, parcel and street geocoding techniques. *Computers, Environment and Urban Systems*, **32**, 214–232.

Zandbergen, P. A. (2009). Geocoding quality and implications for spatial analysis. *Geography Compass*, **3**, 647–680.

Zhao, L., Ochieng, W. Y., Quddus, M. A. and Noland, R. B. (2003). An extended Kalman filter algorithm for integrating GPS and low cost dead reckoning system data for vehicle performance and emissions monitoring. *Journal of Navigation*, **56**, 257–275.

Zhao, Y. (1997). *Vehicle location and navigation systems*. Artech House, Inc.