

Integrating Privacy and Safety Criteria into Planning Tasks

Anna Lavygina, Alessandra Russo, and Naranker Dulay^(✉)

Department of Computing, Imperial College London, London SW7 2RH, UK
{a.lavygina,a.russo,n.dulay}@imperial.ac.uk

Abstract. In this paper we describe a new approach that uses multi-criteria decision making and the analytic hierarchy process (AHP) for integrating privacy and safety criteria into planning tasks. We apply the approach to the journey planning using two criteria: (i) a willingness-to-share-data (WSD) metric to control data disclosure, and (ii) the number of unsatisfied safety preferences (USP) metric to mitigate risky journeys.

Keywords: Personal safety · Information privacy · Multi-criteria decision making · Analytic hierarchy process · Smart city applications

1 Introduction

Smart cities aim to increase the quality of life in cities by addressing problems such as traffic congestion, air pollution, and energy consumption. To help achieve this, information and communication technologies need to be integrated into the infrastructure of the city, to improve its functionality and efficiency [8, 20, 23]. In this paper we consider the provision of transportation services that provide flexible transportation options to efficiently move people around the city. The challenge is to be able to provide individual journeys that are efficient in terms of city-level parameters (e.g. energy consumptions, environmental impact) as well as to satisfy individual preferences and constraints. The latter are typically modelled using a utility function defined over the cost and duration of journeys, and the preferred modes of transportation. However, there can be other considerations, such as, seeking journeys suitable for people with special needs (e.g. the elderly, disabled), avoiding particular areas of the cities (e.g. with a high crime rate, crowded or uncrowded, areas with high pollution levels), not wanting to be tracked (e.g. by video surveillance cameras, MAC address tracking by public WiFi hotspots), not wanting to disclose unnecessary private data (e.g. date of birth to service providers). Such preferences may also be contextual - only apply in particular situations. The aim is to be able to fuse context information with user requirements and then to use the resulting knowledge to make *smart* (automatic or semi-automatic) decisions for users.

Our overall contribution is to demonstrate that the analytic hierarchy process (AHP), a multi-criteria decision making approach, is able to integrate privacy

and safety¹ criteria into planning tasks. We demonstrate this for the classical journey planning task taking into account the traditional utility of the journey plus two additional user-defined criteria, newly introduced in this paper: (i) a novel willingness-to-share-data (WSD) metric that reflects the users perceived sensitivity of their personal data and (ii) the number of unsatisfied safety preferences (USP) which allows users to minimize safety risks.

This paper is organised as follows. Section 2 describes related work. Section 3 outlines our approach to decision making for planning tasks incorporating privacy and safety criteria. Section 4 introduces AHP and shows how to apply it to the journey planning task. The decision making criteria we use (utility, USP, and WSD) are discussed in Sect. 5. This section also provides detailed examples of setting up the criteria and evaluating them for the journey planning task. Section 6 describes a study on how different ratios of the importance of criteria affect the final ranking of journey alternatives. Possible extensions and modifications of the USP and WSD criteria are discussed in Sect. 7. Finally, Sect. 8 concludes the paper and outlines our future work.

2 Related Work

Privacy and Safety in Smart Cities. Research related to privacy and safety of individuals in smart city services has focused on what the city infrastructure, technology and management can do to ensure an individual’s privacy and safety. For example, utilising video surveillance systems to detect and identify abnormal activities, which can help to reduce the level of crime and speed up the response of emergency services [7,21].

In [12] quantitative risk assessment is used to support the design of physical security systems by optimizing the coverage of protection mechanisms. The features that influence fear of crime (e.g. low lighting, desolation, lack of opportunities for surveillance by the general public, etc.), and the ways of reducing the levels of crime and fear of crime (e.g. criminal justice systems, problem-oriented policing, environmental criminology, situational crime prevention) are identified and discussed in [16]. These are important problems for cities to solve. However, they all address the problem “in the large”, by reducing the overall levels of crime and harassment (hence, the risk and fear to become a victim), rather than helping individuals to satisfy their personal safety requirements which may differ from those of the city in kind and/or in degree.

Ferraz and Ferraz [11] identify nine risks associated with information sharing including access to information from applications, information tracking, citizen tracking, and user/citizen data loss. Martinez-Balleste et al. [19] define a citizens’ privacy model based on five dimensions: identity privacy, query privacy, location privacy, footprint privacy and owner privacy. For each dimension they show how existing privacy enhancing techniques (e.g. statistical disclosure control, private information retrieval, privacy-preserving data mining, etc.) can be used

¹ We use term safety to encompass physical security, physiological harm (pollution), physical harm (attack), psychological fear (crowding), etc.

to maintain citizens' privacy. De Cristofaro and Di Pietro [9] focus on query privacy in urban sensing systems. These papers focus on protecting information that is already collected from sensors, users and mobile devices, rather than controlling which data can be sensed, collected, or shared at the first place. The need for usable privacy policies and user interfaces that maximise user control based on their perception of privacy risks is highlighted in [6].

Criteria for Journey Planning. There are a growing number of studies dedicated to understanding travellers' attitudes and criteria for evaluating service quality from a user's perspective (see e.g. [1, 5, 10, 18]). Eboli and Mazzulla [10] report on the importance of service quality indicators, such as reliability, punctuality, pollution, and comfort. In [18] 29 different criteria of users perception of a bus service are grouped into service design, access to service, operation, information and facilities, ticket price and safety. Lynch and Atkins discovered high levels of perceived insecurity for walking at night, in parks and subways and when waiting for public transport services in isolated areas [17]. These studies highlight the importance of user's perceptions of city services. However, none of them show how the considered criteria can be applied in practice, for the evaluation and selection of journeys in a given context. Existing journey planners do not allow any other criteria apart from the preferred modes of transport (e.g. bus, metro), maximum walking time or the need for step free access. Furthermore, they only use these criteria to filter out journey alternatives and rank journeys by travel time only.

The use of utility functions for evaluating and ranking journeys is explored in [2, 15] where different types of utility functions are used based on travel time and cost. Kim et al. [14] proposes a more complex utility function based on various latent variables, such as comfort, convenience, environmental preferences, that is used for building a general choice modelling framework for analysing travellers' choice behaviour rather than planning of an individual journey.

3 Approach

Our approach accounts for both privacy and safety criteria in the decision making process for planning tasks, particularly for journey planning. This approach allows the user to address two crucial aspects: (i) what personal data is sensitive and the degree of data sensitivity, and (ii) what situations are consider as safe/unsafe and which level of safety should be achieved.

The approach can be generalised into any user privacy and safety criteria and any planning task. Our overall approach is shown in Fig. 1. Given a user query and a set of criteria, a set of alternative plans satisfying the query are generated by a planner and passed to ranking process that evaluates all generated plans according to the criteria, for example, the utility of the plan, safety, privacy, reliability, comfort. After ranking, an ordered list of the plans is returned to the user alongside values of metrics that represent the quality of the plans and that can help the user (or user's agent) to select or reject plans.

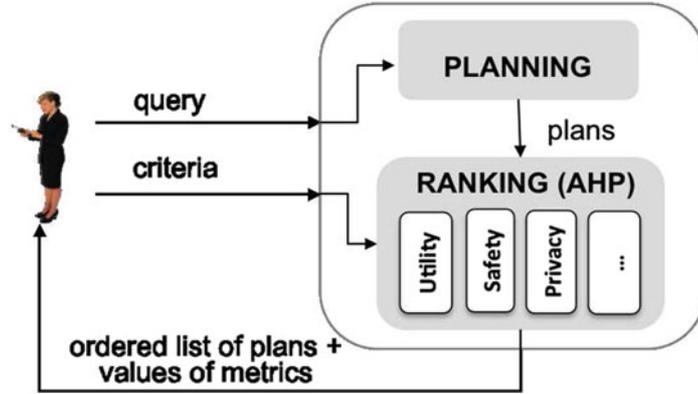


Fig. 1. Multi-criteria planning

As an example, consider a user query “Find the best journey from A to B” and a set of privacy and safety criteria defined by the user. Based on the request, the journey planner generates a list of alternative journeys from A to B that fulfil the user query. We then use AHP to evaluate and rank journey alternatives according to the following three criteria:

- *utility* of a journey based on the time and ticket price (see Sect. 5.1)
- *number of unsatisfied safety preferences* (USP) that allows users to avoid areas with high-crime rates, crowded buses, providers with a bad safety record (see Sect. 5.2)
- *value of a willingness-to-share-data metric* (WSD) that reflects willingness of a customer to share personal information required by a service provider in order to provide a requested service (see Sect. 5.3)

Sections 5.2 and 5.3 also provide the details on how users can define privacy and safety criteria, respectively.

This problem can be tackled using composite objective functions as the weighted sum of all objectives. However, this approach has two major disadvantages. First, solutions are very dependent on the weight-vectors used and in different situations different weight-vectors have to be used [24]. Secondly the values of composite objective functions are often difficult to interpret for complex problems with many criteria. We address these issues by considering all criteria separately and using AHP (see Sect. 4) for ranking alternative solutions. Using AHP in our approach provides (i) means of deriving the weights of the criteria from a series of pairwise comparisons that are more understandable by users - this is important because the overall ranking is dependent on the relative importance of criteria w.r.t. to the goal (ii) tolerance to minor inconsistencies in defining criteria importance, (iii) the ability to deal with both qualitative and quantitative criteria based on either subjective user opinion or actual measurements, (iv) an elegant method to incorporate diverse criteria, such as privacy, safety, comfort level and punctuality.

4 The Analytic Hierarchy Process

The *Analytic Hierarchy Process* (AHP) was introduced by Thomas Saaty [22] and is a multi-criteria decision making approach that allows decisions to be made based on priorities using pairwise comparisons. AHP is widely used in supplier selection [13] and logistics [4]. AHP works as follows: Assume there are n evaluation criteria, and m alternative solutions that have to be ranked according to these criteria. First, weights of criteria are calculated based on pairwise comparisons of the importance of criteria; higher weights correspond to more important criteria. Then, all alternatives are compared pairwise with respect to *each criterion separately*. Finally the results of both series of comparisons are synthesised to give a final ranking of alternatives.

4.1 AHP Hierarchy

The first step in AHP involves decomposition of the problem into a hierarchy of criteria and alternatives. In the AHP hierarchy for planning a journey (see Fig. 2), the goal is to choose the best journey. We consider the following criteria for decision making:

- utility of a journey for the user (to be maximized);
- number of unsatisfied safety preferences (USP) (to be minimized);
- value of the willingness-to-share-data (WSD) metric (to be minimized).

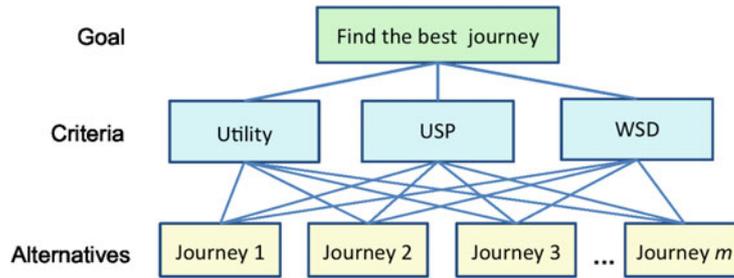


Fig. 2. AHP hierarchy for journey selection.

4.2 Relative Importance of Criteria

To capture the relative importance of criteria, a matrix \mathbf{C} of pairwise comparisons of criteria is created. The matrix $\mathbf{C} = (c_{jk})$ is of dimension $n \times n$, where n is a number of criteria and each element c_{jk} is the relative importance of the j th criterion to the k th criterion with respect to the goal. The elements c_{jk} satisfy the constraint

$$c_{jk} \times c_{kj} = 1, \quad (1)$$

Table 1. Scale of relative importance of criteria

Level of relative importance	Definition
1	Equal importance
3	Moderate importance
5	Essential or strong importance
7	Very strong importance
9	Extreme importance (the highest possible)
2, 4, 6, 8	Intermediate values
1.1, 1.2, 1.3,	Very close importance

where $c_{jk} > 1$ indicates that the j th criterion is more important than the k th criterion. Consequently, in the case where the j th criterion is less important than k th criterion, we have $c_{jk} < 1$, and if the two criteria are indifferent we have $c_{jk} = 1$; which also implies that $c_{jj} = 1$. Saaty [22] suggests a numerical scale between 1 and 9 to express the importance of one criterion over another (see Table 1). Pairwise comparisons can be done by the user when defining criteria for journey planning.

A useful advantage of AHP is that it tolerates minor inconsistencies in the comparisons. For example, assume we have three criteria where criterion #1 is slightly more important than criterion #2, and criterion #2 is slightly more important than criterion #3. If the user asserts that criterion #1 is much more important than criterion #3, then these comparisons are consistent. A minor inconsistency can be induced if the user asserts that criterion #1 is slightly more important than criterion #3; AHP would tolerate this inconsistency. An unacceptable inconsistency would be one where the user asserts that criteria #1 and #3 are indifferent.

Once the criterion importance matrix \mathbf{C} has been established, it can be used to derive the *criteria weight vector* \mathbf{w} using the equation

$$w_j = \frac{\sum_{l=1}^n \bar{c}_{jl}}{n} \quad (2)$$

where $\bar{c}_{jl} = c_{jl} / \sum_{k=1}^n c_{kl}$ is the normalized relative importance.

4.3 Ranking of Alternative Plans

At this step we have to score all generated alternative plans with respect to *each criteria*. To derive these scores we calculate a matrix of pairwise comparisons of alternative plans $\mathbf{B}^j = (b_{ih}^j)$, where b_{ih}^j is the evaluation of the i th alternative plan compared to the h th alternative plan with respect to the j th criterion.

Let x_i^j and x_h^j be the values of the j th criterion for alternative plans i and h respectively.

If the j th criterion has to be maximized, then for all alternative plans i and h with $x_i^j \geq x_h^j$, the element b_{ih}^j can be computed as

$$b_{ih}^j = 8 \frac{x_i^j - x_h^j}{x_{max}^j - x_{min}^j} + 1 \quad (3)$$

where x_{max}^j and x_{min}^j are the maximum and minimum values of the j th criterion.

Similarly, if the j th criterion has to be minimized, then for all alternative plans i and h with $x_i^j \leq x_h^j$, the element b_{ih}^j can be computed as

$$b_{ih}^j = 8 \frac{x_h^j - x_i^j}{x_{max}^j - x_{min}^j} + 1 \quad (4)$$

Similar to the criterion importance matrix \mathbf{C} , the elements of alternative comparison matrix \mathbf{B}^j have to satisfy the constraint $b_{ih}^j \times b_{hi}^j = 1$.

Having obtained \mathbf{B}^j , we can now calculate the score vectors \mathbf{y}^j for alternative plans with respect to each criterion $j \in [1, n]$. This calculation is done using Eq. 2 but replacing the terms c_{jl} with b_{ih}^j .

The score vectors are then used to create a score matrix $\mathbf{Y} = [\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^n]$, from which a plan ranking vector $\mathbf{v} = (v_i), i \in [1, n]$, is calculated by

$$\mathbf{v} = \mathbf{Y} \cdot \mathbf{w} \quad (5)$$

The greater the value v_i , the more preferable the i th alternative plan is.

5 Criteria

We demonstrate application of our approach for the classic journey planning task. In this section we describe three criteria we use to evaluate journey alternatives generated by the journey planner: utility of a journey alternative, the USP and WSD metrics. Detailed examples of setting up the privacy and safety preferences and evaluating the USP and WSD criteria are also included.

5.1 Utility

Journey planners typically rank journeys based on either journey time, walking distance or number of changes. For example, in [2] the authors use a utility function for journey ranking and selection: for each journey i generated by the journey planner, the utility function value is calculated based on the total travel time and ticket price:

$$u_i(t, T_i, c_i) = \frac{e^{(t-T_i)/60}}{e^{|t-T_i|/60+c_i/100}} \quad (6)$$

where $u_i(t, T_i, c_i) \in [0, 1]$ is the utility, t the desired travel time as defined by the user, T_i the travel time of the i th journey, and c_i the total cost of the i th journey. The objective is to find a journey with the highest utility. In this formula if $t - T_i > 0$ (i.e. the traveler arrives in time), then the utility is constant with respect to T_i . If $t - T_i < 0$ (i.e. the traveler arrives late), then the traveler's utility decreases as T_i increases.

In this paper we introduce a penalty for longer journeys even if the traveler is on time, hence our revised utility is:

$$u_i(t, T_i, c_i) = \frac{1}{e^{T_i/t+c_i/100}} \quad (7)$$

Of course, one could devise a more complex utility function. However, we kept the function simple, because AHP uses pairwise comparisons of the journey alternatives with respect to this criteria rather than a very accurate value of the utility function for each alternative plan.

5.2 Unsatisfied Safety Preferences (USP)

Consider the following. Alice is traveling late at night. Suppose that there are two alternative journeys with similar ticket price and travel time, but the first alternative includes walking through an undesirable area. If Alice is aware of this area and concerned about her safety, then she is likely to choose the second option. However, if Alice is not familiar with this area, then she might choose the undesirable area but would have preferred the other alternative.

To address this kind of requirement we propose to include personal safety preferences to other requirements the user can set when starting or changing a journey. For example:

- (i) avoiding areas with high-crime rates, or are sparsely-lit or sparsely-populated,
- (ii) avoiding using trains or buses carrying a low number of passengers for fear of attack,
- (iii) avoiding crowded areas or crowded trains or buses,
- (iv) avoiding service providers with poor safety records or a bad reputation.

All solutions generated by the journey planner have to be evaluated with respect to all safety preferences defined by the user. The *number of unsatisfied safety preferences* (USP) is then used as an *criterion* in the ranking of alternative plans and choice of a journey.

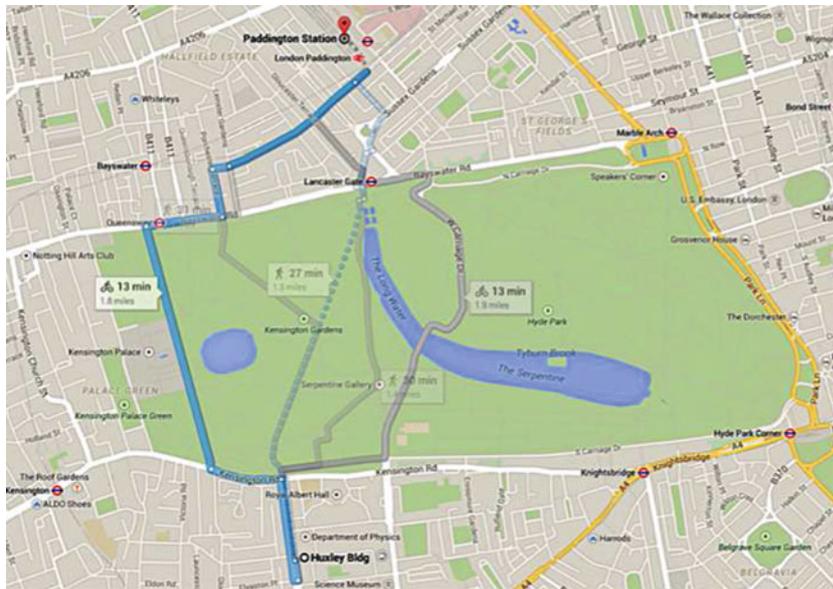
Example 1

Alice arranges a dinner with a friend for 8pm next to Paddington train station. Because she is travelling alone, she wants to avoid crowded areas and crowded transportation as well as sparsely populated areas and transportation. She sends a request to the Journey Planning Service with the following data:

- Starting point: : 180 Queen’s Gate, London SW7 2RH, UK
- Destination point: Paddington Station, Praed St, London W2, UK
- Arrival time: 20:00
- Safety preferences: (i) Avoid sparsely populated areas & transport (ii) Avoid crowded areas & transport

Table 2. List of journey alternatives

Alternative 1 7:28 PM–7:51 PM 23 min, cost £3.80	(1) walk – 5 min – Huxley Bldg to Royal Albert Hall (2) bus 9 – 5 min – Royal Albert Hall to High Street Kensington (3) walk – 3 min – from High Street Kensington to High Street Kensington Underground station (4) the Underground, Circle line – 6 min – High Street Kensington Underground station to Paddington
Alternative 2 7:28 PM–7:57 PM 29 min, cost:£11.50	(1) walk – 2 min – Huxley Bldg to Imperial College Elvaston Pl (2) bus 70 – 17 min – Imperial College Elvaston Pl to Queensway Westbourne Grove (3) walk – 10 min – Queensway Westbourne Grove to Paddington Station
Alternative 3 27 min, cost: £0	(1) walk – 27 min – from Huxley Bldg to Paddington Station via Queens Gate
Alternative 4 30 min, cost: £0	(1) walk – 30 min – from Huxley Bldg to Paddington Station via Queens Gate
Alternative 5 27 min, cost: £0	(1) walk – 27 min – from Huxley Bldg to Paddington Station via Queens Gate
Alternative 6 13 min, cost: £0	(1) cycle – 13 min – from Huxley Bldg to Paddington Station via Broad Walk
Alternative 7 13 min, cost: £0	(1) cycle – 13 min – from Huxley Bldg to Paddington Station via W Carriage
Alternative 8 9 min, cost: £11	(1) taxi – from Huxley Bldg to Paddington Station
Alternative 9 9 min, cost: £7–9	(1) Uber – from Huxley Bldg to Paddington Station

**Fig. 3.** Example of journey alternatives for the journey planning problem in the urban mobility scenario

We simulated the result of this request using Google Maps, taxi services and the Uber application. The generated list of journey alternatives is presented in Table 2.

Alternatives 1, 2, 7, 8 and 9 satisfy both safety preferences. Alternatives 3–6 do not satisfy “Avoiding sparsely populated areas/transport” as all of them are passing through a big park area (Kensington Gardens) as shown in Fig. 3. This area is sparsely populated at the time the request was made because it gets dark early at that time of the year. Hence, alternatives 1, 2, 7, 8 and 9 are preferable to alternatives 3–6 if satisfaction of safety preferences is important for Alice.

5.3 Willingness-to-Share-Data (WSD)

The willingness-to-share-data (WSD) metric is used to control the disclosure of personal data to others, e.g. to service providers. We define the sensitivity of personal data by how much a person values the data in case of a possible harm due to misuse, loss or disclosure by a recipient of the data, such as service provider. We allow different data attributes to be of different sensitivity (for example, a person can define an email address as less sensitive than a phone number or a postal address). Moreover, sensitivity may be dependent on a recipient: the user may trust some recipient to handle their data more than others, and thus may be more willing to provide it.

We define the sensitivity of a certain data attribute as a function $s : (a, p) \rightarrow [0, 1]$, where a is a data attribute and p is a recipient (service provider). Higher values of sensitivity correspond to data that the user prefers not to share.

Given the sensitivity levels for all personal data attributes defined by the user, we define a willingness-to-share-data (WSD) metric that indicates the sensitivity of the whole set of attributes requested by a service provider in order to complete a transaction when ordering a service. We propose the following metric:

$$d_c(\mathbf{a}, \mathbf{r}, p) = \frac{1}{m} \sum_{j=1}^m s(a_j, p) \times r_j \quad (8)$$

where $\mathbf{a} = (a_1, \dots, a_m)$ is a vector of data attributes (e.g. name, address, etc.) that can be possibly requested by a service provider p in order to provide a service, and $s(a_j, p) \in [0, 1]$ is a user-specified level of sensitivity of sharing information related to the j th data attribute with a provider p . The vector $\mathbf{r} = (r_1, \dots, r_m)$ represents the data request mask, and consists of values $r_j = 1$ if the j th data attribute is requested by a provider p , and $r_j = 0$ otherwise.

Users can define the sensitivity of their personal data by completing a form on a mobile application. For this the user has to select a degree of sensitivity ranging from “not sensitive” to “extremely sensitive” (see Table 3 for possible degrees of sensitivity) that are then translated to a value in the range $[0,1]$. For *each* information attribute, the user can assign a configuration of providers using *one* of the following options:

Table 3. Scale of data sensitivity

Sensitivity	Definition
0	not sensitive
0.25	slightly sensitive
0.5	sensitive
0.75	very sensitive
1	extremely sensitive

1. Apply a specified sensitivity level to all service providers.
2. Define *different* levels of sensitivity for two of the following categories of service providers based on a certain level of trust:
 - level of trust greater or equal to x ;
 - level of trust lower than x ;
Levels of trust x are specified in the range $[0,1]$, where higher values correspond to more trustworthy service providers. The trustworthiness of each service provider is calculated based on feedback of all registered users. For the first category of service providers the level of sensitivity must be higher than for the second category.
3. Define a sensitivity level for a specified list of providers (the user has to create the list herself), and set a *different* sensitivity level for other service providers that do not belong to this list.

Note that the WSD metric can be used to express a user’s preferences in the case where *data sharing is negotiable*. If the user does not want to share any data attributes, hard constraints need to be introduced. If a particular journey alternative contains any violation of these constraints, then it is discarded immediately.

Example 2

In this example we calculate the values of the WSD metric s for the scenario described in Example 1. Assume the data attributes $\mathbf{a} = (a_1, \dots, a_m) = \{\text{name, date of birth, email, phone number, postal code, address, GPS location data, payment details}\}$ to be the personal data attributes that could possibly be requested. The following service providers p are available to fulfil Alice’s request: the bus service, the Underground, taxis, and Uber. Suppose, regardless of the service provider p , Alice defines her date of birth, address and phone number as sensitive data, her phone number, GPS location and payment details as very sensitive, and all other attributes she defines as not sensitive. For Table 3 this will yield the following sensitivity levels:

$$\begin{aligned}
s(\text{name}, p) &= s(\text{email}, p) = s(\text{postal code}, p) = 0, \\
s(\text{date of birth}, p) &= s(\text{address}, p) = 0.5 \\
s(\text{phone number}, p) &= s(\text{GPS}, p) = s(\text{payment details}, p) = 0.75
\end{aligned}$$

Some services, such as the bus service and the Underground, do not require any personal data about passengers. For these, we have a data request mask of $\mathbf{r} = (0, 0, 0, 0, 0, 0, 0, 0)$ and a WSD metric of $d_c(\mathbf{a}, \mathbf{r}, \text{Underground}) = d_c(\mathbf{a}, \mathbf{r}, \text{bus}) = 0$.

A taxi service typically requires the phone number to be provided, or $\mathbf{r}^{\text{taxi}} = (0, 0, 0, 1, 0, 0, 0, 0)$. Therefore, $d_c(\mathbf{a}, \mathbf{r}^{\text{taxi}}, \text{taxi}) = (0 \cdot 0 + 0.5 \cdot 0 + 0 \cdot 0 + 0.75 \cdot 1 + 0 \cdot 0 + 0.5 \cdot 0 + 0.5 \cdot 0 + 0.5 \cdot 0) / 8 \approx 0.094$.

Uber requires name, email, the phone number, postal code, payment details and GPS location data to register to their service. Hence, for this service we have $\mathbf{r}^{\text{Uber}} = \{1, 0, 1, 1, 1, 0, 0, 0\}$ and the WSD metric of $d_c(\mathbf{a}, \mathbf{r}^{\text{Uber}}, \text{Uber}) = (0 \cdot 1 + 0.5 \cdot 0 + 0 \cdot 1 + 0.75 \cdot 1 + 0 \cdot 1 + 0 \cdot 0 + 0.5 \cdot 0 + 0.75 \cdot 1 + 0.75 \cdot 1) / 8 \approx 0.281$.

Example 3

Bob lives in a city with smart transportation system that allows people to use the following transportation modes: (i) public transport (public buses, trams, taxis), (ii) FlexiBuses whose routes and stops are determined by passenger requirements, and (iii) car pooling with people sharing a car to save fuel costs and/or gain access to car pooling lanes [3].

Assume the data attributes are the same as for the previous example: $\mathbf{a} = (a_1, \dots, a_m) = \{\text{name, date of birth, email, phone number, postal code, address, GPS location data, payment details}\}$

Bob defines the sensitivity of his data as follows:

- name – not sensitive regardless of the provider: $s(\text{name}, p) = 0$;
- date of birth – sensitive regardless of the provider: $s(\text{date of birth}, p) = 0.5$;
- email – slightly sensitive regardless of the provider: $s(\text{email}, p) = 0.25$;
- phone number – very sensitive for providers a level of trust greater than or equal to 0.8 and extremely sensitive for providers with a level of trust lower than 0.8:

$$s(\text{phone number}, p) = \begin{cases} 0.75, & \text{if } \text{trust}(p) \geq 0.8 \\ 1, & \text{if } \text{trust}(p) < 0.8 \end{cases}$$

- postal code, address and payment details – extremely sensitive regardless of the provider: $s(\text{postal code}, p) = s(\text{address}, p) = s(\text{payment details}, p) = 1$;
- GPS location data – slightly sensitive for the taxi and FlexiBus providers and extremely sensitive for all other providers:

$$s(\text{GPS}, p) = \begin{cases} 0.25, & \text{if } p \in \{\text{taxi}, \text{FlexiBus}\} \\ 1, & \text{otherwise} \end{cases}$$

For this example assume that there are only two providers, FlexiBus (with trust level 0.9) and car pooling (with trust level 0.75) able to fulfill Bob's request. FlexiBus requires name, email, phone number, GPS data, and car pooling provider requires name, phone number and GPS data. The WSD metric values for these two providers are as follows:

For FlexiBus : $d_c(\mathbf{a}, (1, 0, 1, 1, 0, 0, 1, 0), \text{FlexiBus}) = 0 \cdot 1 + 0.5 \cdot 0 + 0.25 \cdot 1 + 0.75 \cdot 1 + 1 \cdot 0 + 1 \cdot 0 + 0.25 \cdot 1 + 1 \cdot 0 = (0.25 + 0.75 + 0.25) / 8 \approx 0.156$

For car pooling: $d_c(\mathbf{a}, (1, 0, 0, 1, 0, 0, 1, 0), \text{car pooling}) = 0 \cdot 1 + 0.5 \cdot 0 + 0.25 \cdot 0 + 1 \cdot 1 + 1 \cdot 0 + 1 \cdot 0 + 1 \cdot 1 + 1 \cdot 0 = (1 + 1)/8 = 0.25$

We can see that although the FlexiBus service is more demanding in terms of the personal data wanted, Bob is more willing to provide his data to this company rather than to the car pooling provider that requires less data but is less trustworthy.

6 The Influence of Criteria Importance Ratios

Based on the scenario described in Examples 1 and 2 we conducted a small study on how different ratios in criteria importance affect the final ranking of journey alternatives. We considered the following cases:

- Case 0:** Only the utility value is used for decision making. For AHP this means that a single criterion, or $n = 1$, is used for ranking journey alternatives.
- Case 1:** Utility is *very much more important* than USP and WSD, which have *equal importance*.
- Case 2:** All criteria are of *equal importance*.
- Case 3:** All criteria are of *equal importance*.
- Case 4:** Utility is *slightly less important* than USP but *equally important* to WSD. USP is *equally important* to WSD (this is a case of moderate inconsistency in criteria ranking).
- Case 5:** Utility is *much less important* than USP and WSD, while USP and WSD are of *equal importance*.
- Case 6:** Utility is *very much less important* than USP and WSD, while USP and WSD are of *equal importance*.

For all cases, the utility values of all alternatives are calculated based on Eq. 7. For alternatives 5 to 8, the USP value is equal to 1, while it is equal to 0 for the other alternatives. The values of the WSD metric were calculated as explained in Example 2. The final vectors of global scores of all journey alternatives and considered cases are presented in Table 4.

We can see from the results that both cycling alternatives (alternatives 6 and 7) have the best (highest) scores when ranking is done based on the utility values only (case 0). This is because they are fast journeys and do not involve any cost. However, when the importance of the USP metric rises (see cases 1 to 6), then the ranking of alternative 6 (also 3 to 5) drops down as it has one unsatisfied preference (which is avoiding an empty area/transport). Similarly, while alternatives 8 and 9 have the same (low) score for case 0 (they are the most expensive alternatives), alternative 8 outperforms alternative 9 as the WSD metric becomes more important. This pattern is due to alternative 8 (taxi) requesting less data from the user to provide a service.

As expected, the scores of alternatives 1 and 2 grow as the importance of utility is decreasing compared to two other criteria. This is due to the fact that these alternatives satisfy the safety preferences specified by the user and do not require any personal information. Nevertheless, these alternatives have lower

Table 4. Global scores of alternatives for the journey planning problem with changing criteria importance

	Global scores of alternatives								
	1	2	3	4	5	6	7	8	9
case0	0.021	0.033	0.123	0.108	0.123	0.248	0.248	0.019	0.019
case1	0.039	0.049	0.112	0.099	0.112	0.214	0.225	0.031	0.029
case2	0.060	0.068	0.098	0.089	0.098	0.173	0.197	0.045	0.041
case3	0.087	0.091	0.081	0.076	0.081	0.123	0.163	0.062	0.055
case4	0.100	0.102	0.069	0.065	0.069	0.097	0.151	0.076	0.069
case5	0.111	0.112	0.066	0.065	0.066	0.078	0.132	0.078	0.068
case6	0.115	0.116	0.064	0.063	0.064	0.070	0.127	0.080	0.070

scores than alternative 7 because utility, yet very low in importance, is still used in the calculation of the final scores.

Interesting results are the final scores of alternatives 4 and 9 for cases 4 to 6. Alternative 9 has a higher final score than alternative 4, although alternative 9 has the worst scores with respect to both the first (utility) and the third (value of the WSD metric) criteria, while alternative 4 is the worst with respect to the second criterion (USP) only. Moreover, the second and the third criteria have the same importance, and feature values of $b_{9,4}^2 = b_{4,9}^3 = 9$ in the pairwise comparison matrices. Nevertheless, the reason that alternative 9 has a higher final score than alternative 4 is due to the way the score vectors of alternatives are calculated (the sum of the scores of all alternatives for each criterion is equal 1), as the difference in scores with respect to the second criteria is greater than first and the third combined.

7 Discussion

Our approach advocates the use of privacy and safety criteria into decision making in planning alongside the common utility of the solutions. Of course, the criteria used for ranking of alternatives can be modified and extended.

The preferences used for calculating the USP metric are related to personal safety. However, there can be various other reasons why a particular user might want to avoid (or not avoid) certain areas or transport. For example, a tourist may want to pass as many places of interest as possible. Similar preferences could also be used for other applications where safety preferences are beneficial, such as hotel booking or choosing a neighbourhood to live in. Moreover, the preferences can be of different importance (for example, for a particular commuter avoiding unsafe areas is more important than avoiding crowded areas). In these cases the weighted sum can be used instead of number of unsatisfied preferences. By using AHP for ranking of alternative solutions there is no need for normalization of the criteria, because all alternatives are compared to each other with respect to

each criteria separately (see Eqs. 3 and 4). Learning user preferences and their relative importance based on the decisions (the final selections) made by a user can further improve the quality of ranking alternatives where there are hidden or context-dependant preferences. We can also organise criteria into a hierarchy to reduce the number of pairwise comparisons.

One can also think about an alternative to the WSD metric to control the information shared with service providers. WSD (see Eq. 8) reflects an “average harm” of sharing all requested by a provider data and might not be effective in cases when providers request not many, but very sensitive data attributes. Using $\max_j s(a_j, p) \times r_j$ in such situations would help to protect the most sensitive data attributes by giving in the less sensitive attributes.

8 Conclusion and Future Work

In this paper we proposed an approach for directly incorporating the privacy and safety criteria into decision making in planning. The approach was illustrated using the classic journey planning task. Our approach allows a user to define their own criteria and their relative importance. AHP was used to rank solutions incorporating two criteria, the number of unsatisfied safety preferences (USP) and a willingness-to-share-data (WSD) metric, plus a utility. The combination of these criteria helps users to find the safer journeys and to control the information they share with providers as well as achieve the required utility. Applying AHP allows to produce not only the ranked list of alternative plans, but also scores for those alternative plans, which can help users to understand why some alternatives are preferable to others, and in some cases select the alternative not from the top of the list.

To conduct some user trials we plan to develop a mobile phone application that combines the approach and criteria we described in this paper with existing journey planning services (e.g. The Google Directions API). For this integration we need to define parameters and contextual data associated with journey alternatives that can be used to calculate USP metric, in particular: assign safety levels to areas and routes, define context-dependent sparseness/crowdedness of areas and routes.

Acknowledgements. This work is supported by the 7th Framework EU-FET project ALLOW Ensembles (grant 600792).

References

1. André, P., et al.: Journey planning based on user needs. In: CHI 2007 Extended Abstracts on Human Factors in Computing Systems, pp. 2025–2030. ACM (2007)
2. Andrikopoulos, V., et al.: A game theoretic approach for managing multi-modal urban mobility systems. In: Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics (AHFE 2014). CRC Press/Taylor & Francis, Kraków, Poland, July 2014

3. Andrikopoulos, V., Bucchiarone, A., Gómez Sáez, S., Karastoyanova, D., Mezzina, C.A.: Towards modeling and execution of collective adaptive systems. In: Lomuscio, A.R., Nepal, S., Patrizi, F., Benatallah, B., Brandić, I. (eds.) ICSOC 2013. LNCS, vol. 8377, pp. 69–81. Springer, Heidelberg (2014)
4. Barker, T.J., Zabinsky, Z.B.: A multicriteria decision making model for reverse logistics using analytical hierarchy process. *Omega* **39**(5), 558–573 (2011)
5. Beirão, G., Cabral, J.S.: Understanding attitudes towards public transport and private car: a qualitative study. *Transp. Policy* **14**(6), 478–489 (2007)
6. Breaux, T.: Privacy requirements in an age of increased sharing. *IEEE Softw.* **31**(5), 24–27 (2014)
7. Calavia, L.: A semantic autonomous video surveillance system for dense camera networks in smart cities. *Sens.* **12**(8), 10407–10429 (2012)
8. Caragliu, A., et al.: Smart cities in europe. *J. Urban Technol.* **18**(2), 65–82 (2011)
9. De Cristofaro, E., Di Pietro, R.: Adversaries and countermeasures in privacy-enhanced urban sensing systems. *IEEE Syst. J.* **7**(2), 311–322 (2013)
10. Eboli, L., Mazzulla, G.: A methodology for evaluating transit service quality based on subjective and objective measures from the passengers point of view. *Transp. Policy* **18**(1), 172–181 (2011)
11. Ferraz, F.S., Ferraz, C.A.G.: Smart city security issues: depicting information security issues in the role of an urban environment. In: 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing, pp. 842–847. IEEE (2014)
12. Flammini, F., Gaglione, A., Mazzocca, N., Pragliola, C.: Quantitative Security risk assessment and management for railway transportation infrastructures. In: Setola, R., Geretshuber, S. (eds.) CRITIS 2008. LNCS, vol. 5508, pp. 180–189. Springer, Heidelberg (2009)
13. Ho, W., et al.: Multi-criteria decision making approaches for supplier evaluation and selection: a literature review. *Eur. J. Oper. Res.* **202**(1), 16–24 (2010)
14. Kim, J., et al.: Hybrid choice models: principles and recent progress incorporating social influence and nonlinear utility functions. *Procedia Environ. Sci.* **22**, 20–34 (2014)
15. Koppelman, F.S.: Non-linear utility functions in models of travel choice behavior. *Transp.* **10**(2), 127–146 (1981)
16. Loukaitou-Sideris, A., Eck, J.E.: Crime prevention and active living. *Am. J. Health Promot.* **21**(4s), 380–389 (2007)
17. Lynch, G., Atkins, S.: The influence of personal security fears on women’s travel patterns. *Transp.* **15**(3), 257–277 (1988)
18. Mahmoud, M., Hine, J.: Using AHP to measure the perception gap between current and potential users of bus services. *Transp. Plann. Technol.* **36**(1), 4–23 (2013)
19. Martinez-Balleste, A., et al.: The pursuit of citizens’ privacy: a privacy-aware smart city is possible. *IEEE Commun. Mag.* **51**(6), 136–141 (2013)
20. Nam, T., Pardo, T.A.: Conceptualizing smart city with dimensions of technology, people, and institutions. In: Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times, pp. 282–291. ACM (2011)
21. Patsakis, C., Solanas, A.: Privacy-aware event data recorders: cryptography meets the automotive industry again. *IEEE Commun. Mag.* **51**(12), 122–128 (2013)
22. Saaty, T.L.: What is the analytic hierarchy process? In: Mitra, G., Greenberg, H.J., Lootsma, F.A., Rijkaert, M.J., Zimmermann, H.J. (eds.) *Mathematical Models for Decision Support*. NATO ASI Series, vol. 48, pp. 109–121. Springer, Heidelberg (1988)

23. Schaffers, H., et al.: Smart cities and the future internet: towards cooperation frameworks for open innovation. In: Domingue, J., et al. (eds.) FI. LNCS, vol. 6656, pp. 431–446. Springer, Heidelberg (2011)
24. Srinivas, N., Deb, K.: Multiobjective optimization using nondominated sorting in genetic algorithms. *Evol. Comput.* **2**(3), 221–248 (1994)