Using Extracted Behavioral Features to Improve Privacy for Shared Route Tracks

Mads Schaarup Andersen, Mikkel Baun Kjærgaard, and Kaj Grønbæk

Department of Computer Science
Aarhus University
{masa,mikkelbk,kgronbak}@cs.au.dk

Summary. Track-based services, such as road pricing, usage-based insurance, and sports trackers, require users to share entire tracks of locations, however this may seriously violate users' privacy. Existing privacy methods suffer from the fact that they degrade service quality when adding privacy. In this paper, we present the concept of *privacy by substitution* that addresses the problem without degrading service quality by substituting location tracks with less privacy invasive behavioral data extracted from raw tracks of location data or other sensing data. We explore this concept by designing and implementing *TracM*, a track-based community service for runners to share and compare their running performance. We show how such a service can be implemented by substituting location tracks with less privacy invasive behavioral data. Furthermore, we discuss the lessons learned from building TracM and discuss the application of the concept to other types of track-based services.

Key words: Location, Privacy, Track-based services, Privacy-By-Substitution, Behavioral Features, Running

1 Introduction

Recently, new types of Location-Based Services (LBSs) have emerged where the foundation of the service is a track - a time-ordered sequence of locations - rather than a single location. These services are called *track-based services* and include application domains, such as, ride-sharing, road-pricing, usage-based car insurance and sport trackers [5].

Ever since location technology appeared on the mass market in special purpose devices and mobile phones, the issue of location privacy has been raised [8]. For track-based services the problem is even more pertinent to address, as several kinds of personal information can be inferred from location tracks [8]. For instance, this raises issues with regards to citizen surveillance in connection with government-based road-pricing or customer surveillance for usage-based car insurances.

A recent survey of methods for location privacy identifies a general lack of methods for track-based services as most existing obfuscation and anonymity methods only consider point-of-interest services [6]. This lack of methods is also

noted by Ruppel et. al [12], who present some of the first attempts to obfuscate a track of locations. A study by Krumm [8] demonstrates the potential hazards in sharing large amounts of location data. Furthermore, three general privacy methods for track-based services are presented, but the methods suffer from potentially degrading service quality. Mun et. al [11] present the privacy method of selective hiding, but this method is only applicable to a limited set of track-based services. Furthermore, existing software infrastructures for supporting track-based services do not even address privacy explicitly [5, 9].

In this paper we present the concept of *privacy by substitution* that addresses the problem without degrading service quality by substituting location tracks with less privacy invasive behavioral data extracted from raw tracks of location data or other sensing data. The behavioral data is then used instead of the location tracks to realise the intended application logic. Furthermore, by extracting the behavioral feature data on users' devices only less invasive data needs to be shared with external parties. We argue for that many possibilities exist for substituting location-tracks with less privacy invasive behavioral feature data to address the individual privacy needs of track-based services [13]. For example, in a running scenario we can extract behavioral features, such as height and pace curves, and length and completion times, which reveals much less information than a time-ordered sequence of locations. In a usage-based car insurance scenario one could extract features, such as, acceleration or deceleration patterns, as well as kilometers where speed limits are exceeded. It depends on the application scenario what data is relevant. In the following we list examples of highly invasive data one should avoid sharing and examples of less invasive data.

Highly Invasive: Location tracks, home or work address, own or family members identity, daily temporal patterns, social values.

Less Invasive: Altitude, pace, speed, bearing, mode of transportation, acceleration profile, accumulated road usage.

The contributions of this paper are as follows: (i) we present the concept of *privacy by substitution*; (ii) We explore this concept by designing and implementing *TracM*, a track-based community service for kids and youngsters to share and compare running performance to promote healthy behavior. We show how the service can be implemented substituting location tracks with less privacy invasive Decorated Height Curves (DHCs) and using similarity comparison techniques to realise the intended application logic; (iii) We present evaluation results for both simulated and real world running tracks that provide evidence that these techniques can compare a runner's performance and identify relevant runners / tracks to virtually compete against. The results indicate that a similarity technique based on normalized Euclidean distances gives the best comparison performance; (iv) Furthermore, we discuss the lessons learned from building TracM and discuss the application of the concept to other types of track-based services.

2 Using Less Invasive Behavioral Data

We explore a community service for runners with the intent to build a smartphone application for kids and youngsters to promote healthy behavior. The intended application is going to be launched by a national agency and therefore a requirement is that it provides protection of the kids' and youngsters' privacy. On the other hand it is known that social community aspects of smartphone applications can provide a strong motivational drive for behavioral change [10]. However, current methods for implementing such social community services require that complete location tracks are shared with external services which might be a problem, e.g., if a stalker can infer the location of a victim at a specific time of day. Therefore, there is a need to apply our concept in this scenario to substitute the sharing of location tracks with less privacy invasive behavioral data to improve the privacy protection while providing community driven functionality.

The three steps involved in applying our concept of privacy by substitution are as follows:

(i) Analyze the service's data requirements and functionality. (ii) Identify a minimal set of behavioral data that can fulfill the data needs of the service. E.g., for the running scenario we identified that decorated height curves can fulfill this need. (iii) Find means to implement the service functionality using behavioral data. E.g, in the running case use similarity techniques to compare decorated height curves and thereby realise the intended application logic.

For step one we have analysed existing track-based community services for runners [2, 3] and identified three types of functionality (F1-F3) to support:

F1 - Share Runners should be able to share their running performance results via social media, e.g., total distance and completion time.
F2 - Compete Runners should be able to compete againts each other.
F3 - Inspire Runners should be able to share tracks to inspire other runners to run new routes.

We will in the following sections cover step two and three and show how to implement **F2** and outline solutions for **F1** and **F3**, using behavioral features.

3 TracM - Behavioral Feature Extraction Services

Building on the previous analysis, we present TracM, a privacy preserving distributed service for implementing track-based social community services for runners. TracM provides privacy preserving implementations of functionality **F1-F3**. Smartphone applications, such as, the mentioned application for kids and youngsters can then be implemented using TracM by implementing a graphical user interface that utilize the TracM functionality. The two main techniques to implement the functionality are Decorated Height Curves (DHCs) and similarity techniques for comparing DHCs. Hence in relation to the principle of privacy by substitution, DHCs and comparison of these become the tool that facilitates the substitution.

Decorated Height Curves consist of tuples of time, length, height and pace values. To compute a DHC TracM collects a regular time-stamped GPS location track which is then transformed into a DHC. A GPS track is transformed into a DHC in the following way:

$$(timestamp, latitude, longitude, altitude) \Rightarrow (time, length, height, pace)$$

The individual values in the tuples are calculated in the following way:

Time time since the user started running to add temporal privacy and therefore $time_0 = 0.0s$

Length distance moved since the user started running and is determined by the distance between consecutive GPS positions starting with $length_0 = 0.0m$.

Height normalized height $h_r^{'}$ of a GPS position, offset by the mean height h_{mean} over the whole location track. For the N height entries of a DHC each entry h_r is normalized as follows:

$$h_r^{'} = h_r - \bar{a} \quad \wedge \quad \bar{a} = \frac{\sum_i^N h_i}{N}$$

Using the mean of maximum and minimum heights for normalization was also considered, but this created problems with erroneous height measurements as they could offset the curve making it different from similar curves with no errors.

Pace in meters per seconds are calculated from pairs of consecutive GPS positions.

The flow in TracM is as follows focusing on **F2: Compete**: (1) Initially the user either selects/creates a route in their own local collection or selects an inspiration height curve (HC) receiving good ranks by other runners provided by a remote DHC repository (**F3: Inspire**). In the later case the service will find the most similar local route, if it exists, matching the height profile and show it to the user. (2) The HC of the route is sent to a comparator which uses a similarity measure to find the most similar HC from the remote DHC repository. (3) This DHC is then sent to the local TracM service. The TracM service continually checks that the user is actually running along the route while informing the user of progress in relation to the DHC.

Track creation/selection and recording of the users track are both done on the TracM device and, hence, what is made publically available is only the HC in (2). Afterwards the user can share summary statistics including the height curve over relevant social media channels (**F1: Share**).

Characteristics of Similar Height Curves A central element, in the above solutions, is to be able to compare HCs. For the comparison we define that two curves' similarity depend on the number of similarity criteria given below that they satisfy:

C1 nearly the same total ascent and total descent.

C2 peaks appear on similar places on the length axis. I.e. in a visual inspection of the curves they should peak at similar times on the length axis.
C3 close to similar minimum and maximum heights.
C4 Two curves which are similar in every aspect besides being shifted along the length axis, should have a high degree of similarity. This is to insure that if two users run the same track, and starts tracking with 20m between the start points they should still be detected as running on a similar track.
C5 Curves should be of similar length within a percentage threshold of $\Theta = \pm 10\%$.
C6 Two curves that only differ in being shifted on the height axis should be exactly similar.

3.1 Similarity Measures

The other main concept in TracM is the similarity between features based on DHCs. Two of the similarity measures we consider in this work are known from shape similarity [14] and the last three from statistics. From shape similarity the following similarity measures were implemented: *Euclidean Distance* and *Integral*. From statistics the following measures were implemented: *Cosine Coefficient*, *Histogram Intersection*, and *Kolmogorov-Smirnov*. For the three latter to make sense, one can think of the height values as the sample set. This requires that the height measurements are spaced equally apart on the length axis. Due to variations in running speed, positioning errors and sample jitter, DHCs from GPS tracks will not be evenly spaced. Therefore we process the DHCs using interpolation to have a height measurement for each twenty meters. It is a problem that some of the similarity measures require curves to have the same length. To adress this we extend the shortest curve by the length missing at the same height as the last measured height. Another option would be to just compare the curves until the shortest is finished, but this was rejected as it would have a significant impact if the route ends in a steep incline.

Euclidean Distance Measure The Euclidean distance measure between two curves C and D is calculated as the sum of the Euclidean distance between the individual points c_i and d_i for i from 1 to N. We expand this measure by taking the K nearest points on the target curve into consideration and take the minimum distance. Furthermore, the result is normalized by the mean distance to be comparable for curves of different length:

$$E(K, C, D) = \frac{\sum_i^N min_k(\sqrt{(c_{k_x} - d_{k_x})^2 + (c_{k_x} - d_{k_y})^2})}{N}$$

Where min_k iterates from $i - K$ to $i + K$ and returns the minimum distance. The measure is referred to as $E(K)$.

Integral Measure The Integral (Int) measure computes the area of symmetric difference of the area spanned by the two curves C and D and negative infinity, defined as

$$I(C, D) = area((C - D) \cup (D - C))$$

The output will be a positive number and the smaller it is, the more similar the curves are.

Cosine Coefficient The Cosine (Cos) similarity measure captures the similarity between two vectors C and D by measuring the angle between them. It examines whether these point in relatively the same direction. In our case the vectors contain equally spaced height entries. The measure is calculated using the following formula:

$$cos(\theta) = \frac{C \cdot D}{\parallel C \parallel \parallel D \parallel}$$

Histogram Intersection Histogram intersection (His) measures the distance between two histograms and is often used as a similarity measure for images.

$$H(C, D) = 1 - \frac{\sum_i min(c_i, d_i)}{\sum_i d_i}$$

Here C and D are the two sample sets of heights represented as histograms. The output is a number between zero and one with one denoting exactly similar curves.

Kolmogorov-Smirnov Distance The Kolmogorov-Smirnov distance (KS), measures the similarity of two sample sets of heights C and D. It is defined as follows for each sample c_i and d_i:

$$K(C, D) = max_i|c_i - d_i|$$

4 Evaluation of Decorated Height Curve Similarity

To find a good similarity measure for DHCs to use for realising the intended application logic, an evaluation framework was developed to test TracM with each of the five aforementioned similarity measures. We consider two versions of the Euclidean metric with K equal to 1 and 10 named E(1) and E(10), respectively. It would be relevant in future work to consider other parameterisations of this metric.

For the evaluation we establish a ground truth using the characteristics from Section 3. To test $C1$, $C3$ and $C5$ statistics for the curves can be computed and compared. $C2$, $C4$ and $C6$ can be tested using visual inspection.

The evaluation is based on simulated as well as real world data. The length of the tracks is selected to be 5 km since the domain is running and 5 km is a distance most people feel comfortable running. According to $C5$, this gives us with our choice of Θ a range from 4.5 to 5.5 km in the real world data of tracks with should be considered similar.

4.1 Simulated Height Curves

The simulated curves are chosen to exercise most of the characteristics from Section 3. The basic set of simulated curves are listed in Table 1. All curves have a height difference of 100 m (except for the flat curve). Therefore $C3$ is always satisfied in the simulation cases.

Table 1. List of simulated curves

$flat$	A flat curve.
sin, sin^{-1}	A sine curve and it's inverse.
asc, des	A curve that evenly ascents from -50 m to 50 m during the entire track and it's inverse.
$peak$, val	A curve that evenly ascents from -50 m to 50 m and halfway descents from 50 m to -50 m evenly and it's inverse.
$peaks$	A curve that evenly ascents from -50 m to 50 m in 250 m and descents from -50 m to 50 m in 250 m. This pattern is repeated throughout the 5 km.

Some of the curves in addition have two shifted versions where they either start out with 500 m in the same height or end in 500 m in the same height (e.g. $peaks_b$ and sin_a). This is specifically to test the criteria $C4$.

For the evaluation using the simulated curves we select three test cases having very different shapes: $peak$, asc and $peaks$. The test cases are named by the source height curve:

Test Case: peak (S1) $peak$ shares total ascent and descent with sin, sin^{-1}, val, $peak_b$, $peak_a$ and, therefore, $C3$ is satisfied. In relation to $C4$ $peak_b$ and $peak_a$ should turn out similar.

It is expected that the most similar curves are the shifted versions of $peak$ followed by sin and it's shifted versions. This is based on $C1$ and $C3$ and the fact that $peak$ and sin peak on the same place on the length axis ($C2$). The top 5 results of each similarity measure can be found in the table of Figure 1 in ascending order.

We notice that all measures agree that shifted versions of $peak$, and sin and it's shifted versions are the most similar as expected. The measures agree on similar curves, but disagree on order. In the figure we also see $peak$, $peak_{after}$, sin, and $flat$ visually. In relation to $C2$ the visual inspection indicate similarity of sin and $peak_{after}$ as they have similar peaks. However, $flat$, which was rated similar by the KS measure, is very dissimilar. Hence, KS is not a good measure in this case.

Test Case: ascent (S2) asc is characterized by having no descent, sharing that feature with $flat$ and sharing total descent with all other curves except $peaks$, $flat$ and des. Results can be found in Figure 2.

Cos	His	KS	Int	E(1)	E(10)
$peak_a$	sin	sin	$peak_a$	$peak_a$	$peak_a$
sin	$peak_a$	$peak_a$	sin	sin	sin
sin_a	sin_a	sin_a	sin_a	sin_a	sin_a
$peak_b$	sin_b	$flat$	$peak_b$	$peak_b$	$peaks_a$
sin_b	$peak_b$	$peaks$	sin_b	sin_b	$peaks$

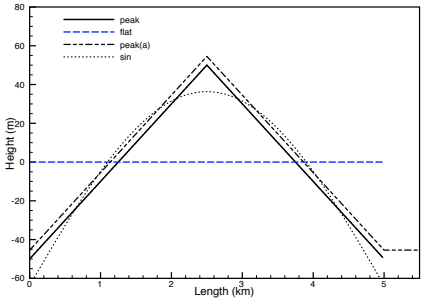

Fig. 1. Table listing most similar curves to *peak* found by the similarity measures, and a subset of these shown visually.

Cos	His	KS	Int	E(1)	E(10)
asc_a	asc_a	asc_a	asc_a	asc_a	asc_a
asc_b	asc_b	asc_b	asc_b	asc_b	asc_b
sin_b	sin_b	$peaks_b$	$flat$	$flat$	$peaks_b$
$peak_b$	$peak_b$	sin_b	$peak_b$	$peak_b$	$peaks$
$peaks_b$	$peaks_b$	$flat$	$peak$	$peaks_b$	$peaks_a$

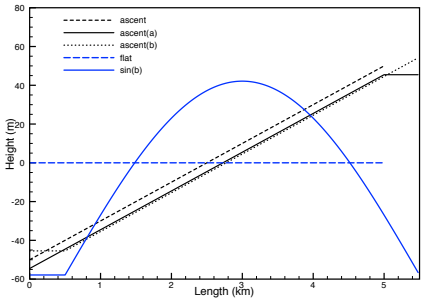

Fig. 2. Table listing most similar curves to *asc* found by the similarity measures, and a subset of these shown visually.

All measures agree on the two most similar curves being the shifted versions of *asc*. But besides from that the measures do not agree. In the figure the shifted versions of *asc* along with two of the other similar curves are shown. The visual inspection indicates that neither sin_b nor $flat$ are similar to *asc*, and hence, the likely reason that the measures disagree is that there are no more than two similar curves. This is confirmed by the output from the measures, where there is a large gap in the values from the shifted curves of *asc* to the next.

Test Case: peaks (S3) *peaks* is characterized by the fact that shifting the curve by 500 m on the length axis has the consequence of producing a curve that is close to the inverse of the original while sharing total ascent/descent (C1), and it should be similar in relation to $C4$. The shifted versions of the curve should be similar in relation to $C2$. Results can be found in Figure 3.

All measures but KS agree that $peaks_a$ is the most similar, but from there on they differ a lot. However, three agree on $flat$ as being the second most similar. In Figure 3, *peaks* and it's shifted versions along with $flat$ and sin_a are shown. *peaks* and $peaks_{before}$ clearly demonstrate the issue described in $C4$. Hence, this curve has to be rated as similar. Only KS and E(10) rate $peaks_{before}$ among the top 5 similar curves and since KS rated $flat$ as the most similar, only E(10) performs adequate according to $C4$.

Cos	His	KS	Int	E(1)	E(10)
$peaks_a$	$peaks_a$	$flat$	$peaks_a$	$peaks_a$	$peaks_a$
sin_a	$flat$	$peaks_a$	$flat$	$flat$	$peaks_b$
$peak_a$	sin	$peak$	val	$peak_a$	$peak_a$
des_b	sin_a	$peaks_b$	asc	des_b	$flat$
des_a	des_b	sin	$peak$	des_a	$peak_b$

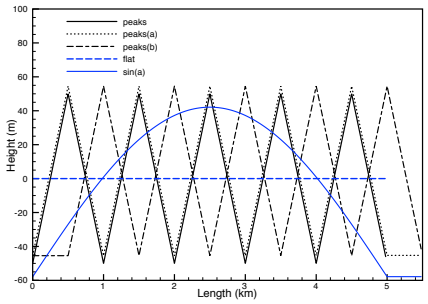

Fig. 3. Table listing most similar curves to *peaks* found by the similarity measures, and a subset of these shown visually.

Overall, the more complex the curves become, the more different the measures perform. For S1 and S2 several measures solved the problem well, but in S3 only E(10) performed as expected. Furthermore, KS and His proved to perform significantly worse than the other measures and, hence, we will leave out their results in the following section as they proved to perform bad on real world data as well.

4.2 Real World Data

To gather data for the real world evaluation we use GPSies.com, a large database of GPS tracks, [4]. As TracM enables users to compete against users from other regions than their own, data from three countries is chosen: Germany, Denmark and The Netherlands. A query for 200 tracks was issued for each of these three countries, and from these 31 tracks were selected at random for comparison (labeled $t1$ to $t31$). In ten repetitions an input curve were selected, leaving a set of curves for comparison of size 30. To establish a ground truth the curves were manually compared by visual inspection to the remaining 30 curves with regards to the criteria $C1$-$C6$. The curves satisfying $C1$-$C6$ were marked as similar and therefore should be identified as similar by the evaluated metrics.

In the following we will discuss one of the test cases and list top 10 most similar curves for each measure. The considered curve is almost flat. Total ascent is 20 m, descent is 20 m, and the height difference is 4 m. Ground truth and result of the test case can be found in Table 2. Here we notice that Cos and E(10) perform good and Int and E(1) worse. Figure 4 shows the input curve, t30, in relation to a very similar curve, t15, and a dissimilar, t28.

A summary of the results of all 10 test curves can be found in Table 3. Here we see accuracy of the different measures for top 5 and top 10 with respect to the ground truth set identified. As the results show, once again E(10) outperforms the other measures. However, even with this measure, some curves are classified incorrectly.

Ground Truth
t3, t4, t14, t15, t20, t21, t26, t31

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Cos	(t7)	t26	t20	t14	t15	(t24)	t4	t21	t3	(t11)
Int	t14	t4	t15	(t24)	(t13)	(t10)	(t12)	(t25)	(t8)	(t22)
E(1)	t14	t4	t15	(t24)	(t8)	(t12)	(t13)	(t10)	(t25)	t26
E(10)	t21	(t11)	t31	t4	t15	t14	(t19)	t26	t20	t3

Table 2. Ground truth and result of one repetition, with top 10 most similar curves for each measure. Numbers in parenthesis are curves not present in ground truth.

Table 3. The average performance, taken over the ten tested curves, of the similarity measures by how many of the ground truth curves were found in the top 5 and top 10 results, respectively.

Test	**Cos**	**Int**	**E(1)**	**E(10)**
Top 5	64.5%	68.3%	68.3%	74.3%
Top 10	73.9%	72.3%	74.0%	91.1%

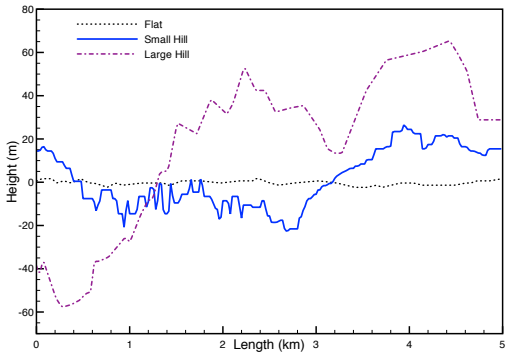

Fig. 4. Three curves from the real world test case.

5 Performance and Potentials Beyond Running Tracks

We presented the concept of privacy by substitution. We evaluated this concept by developing, implementing, and evaluating TracM and showed that it is indeed possible to find less privacy invasive behavioral data to provide the functionality of a track-based community service for runners. I.e. with the current results it is indeed possible to create privacy enhanced services for kids and youngsters.

A limitation of our current implementation of TracM is that we did pre-processing when comparing two curves of unequal length, by choosing a naive strategy of prolonging the shorter curve with a flat piece based on last height. It would be relevant to explore how changing this strategy would effect the similarity. The measure E(10) had the best performance eventhough not flawless. Therefore as an additional element one could let the user check the proposed curve by visual inspection and select another if unsatisfied.

The advantage of the approach used in TracM, is that it is a concept that can be applied to a range of services. However, this also has the disadvantage that the method does not directly prescribe how to solve the privacy problem in

a particular application as the choice of behavioral data and similarity measures have to meet the specific requirements of that application. To generalize the method, we will apply the method to several domain cases and hope to develop a toolkit to support developers in providing privacy based on feature extraction techniques.

As we illustrated with TracM it takes some effort to determine the behavioral data that should replace tracks. This is mainly due the fact that this was the first application domain to which the concept was applied, but also because it does require a different mindset overall. The case of community based track sharing for runners is, however, relatively easy compared to other domains as no large corporations or government agencies have to base their business on the functionality as they would have to in usage-based insurance or road pricing scenarios.

To explore the generally applicability for other application domains in track-based services, let us briefly examine usage-based insurance, of which the Alka Box [1] is an example. Currently, such systems are based on location tracks of the user, but instead of calculating insurance premium based on *where* the user has driven, it might make more sense to base it on *how* he drives. This is based on the assumption that, in car insurance, an aggressive driving style is more likely to capture how likely a user is to be in an accident rather than where he drives. The driving style might be estimated by analyzing features such as the acceleration/deceleration patterns in relation to speed, as well as the amount of kilometers where speed limits are exceeded, etc.. This leads to pattern matching and hence similarity measures can be used to solve the problem.

This indicates that this approach can also be used for other application domains in track-based services. However, actual implementations of such systems are needed to further evaluate the potential of the concept.

6 Conclusion

In this paper, we introduced the concept of privacy by substitution where less privacy invasive behavioral feature data is shared instead of complete location tracks to improve privacy. To apply the concept one has to identify the least amount of behavioral data needed to enable a specific track-based service and techniques for using the behavioral data to realise the intended application logic. We applied this concept to the domain of community-based running track sharing and design and implemented TracM, a service supporting feature extraction based on decorated height curves and similarity measures. Furthermore, we evaluated five similarity measures with TracM on simulated and real world data, and found that a normalized Euclidean distance had the best similarity performance. Furthermore, we argued that the concept has a more general applicability exemplified by usage-based insurance and road-pricing.

In our ongoing work we are trying to address the following: First, we will deploy the TracM service in the context of a mobile application to study whether users feel that it provides a similar service to existing services and if such a service

can be efficiently implemented [7]. Second, we propose to explore the presented concept for other track-based services with emphasis on road-pricing and usage-based insurance. Finally, we propose to implement a wider array of similarity measures which can be used in adding privacy to track-based services.

Acknowledgments

This work is supported by a grant from the Danish Council for Strategic Research for the project: EcoSense.

References

1. Alka box, 2012. http://www.alkabox.dk/.
2. Endomondo, 2012. http://www.endomodo.com/.
3. Garmin connect, 2012. http://connect.garmin.com/.
4. Gpsies.com, 2012. http://www.gpsies.com/.
5. Ganesh Ananthanarayanan, Maya Haridasan, Iqbal Mohomed, Doug Terry, and Chandramohan A. Thekkath. Startrack: a framework for enabling track-based applications. In *Proc. of the 7th Int. Conf. on Mobile Systems, Applications, and Services*, 2009.
6. M. S. Andersen and M. B. Kjærgaard. Towards a new classificaion of location privacy methods in pervasive computing. In *Proc. of the 8th Int. ICST Conf. on Mobile and Ubiquitous Systems*. Mobiquitous, 2011.
7. Mikkel Baun Kjærgaard. Location-based services on mobile phones: Minimizing power consumption. *IEEE Pervasive Computing*, 11(1):67–73, 2012.
8. John Krumm. Inference attacks on location tracks. In *Proceedings of the 5th international conference on Pervasive computing*. Springer-Verlag, 2007.
9. J. Langdal, K. Schougaard, M. Kjærgaard, and T. Toftkjær. Perpos: a translucent positioning middleware supporting adaptation of internal positioning processes. *Middleware 2010*, pages 232–251, 2010.
10. Florian 'Floyd' Mueller and Stefan Agamanolis. Sports over a distance. *Comput. Entertain.*, 3, 2005.
11. Min Mun, Sasank Reddy, Katie Shilton, Nathan Yau, Jeff Burke, Deborah Estrin, Mark Hansen, Eric Howard, Ruth West, and Péter Boda. Peir, the personal environmental impact report, as a platform for participatory sensing systems research. In *Proc. of the 7th int. conf. on Mobile systems, applications, and services*. ACM, 2009.
12. Peter Ruppel, Georg Treu, Axel Küpper, and Claudia Linnhoff-Popien. Anonymous user tracking for location-based community services. In *Proc. of. 2nd Int. Workshop on Location- and Context-Awareness*. Springer, 2006.
13. Marcello Paolo Scipioni and Marc Langheinrich. I'm here! privacy challenges in mobile location sharing. *2nd Int. Workshop on Security and Privacy in Spontaneous Interaction and Mobile Phone Use (IWSSI/SPMU 2010)*, 2010.
14. R.C. Veltkamp. Shape matching: similarity measures and algorithms. In *Shape Modeling and Applications, SMI 2001 International Conference on.*, pages 188 – 197, may 2001.