MatMap: An OpenSource Indoor Localization System

Richard Ižip and Marek Šuppa

Faculty of Mathematics, Physics and Informatics,
Comenius University, Bratislava, Slovakia
izip1@uniba.sk, suppa1@uniba.sk,

Abstract. In the past few decades a huge amount of work has been done in the area of Indoor Localization Techniques and Systems.
In this paper we present MatMap: A simple OpenSource Indoor Localization System which can be used on any device that is capable of scanning WiFi access points in its proximity, does not require addition of any specialized hardware in the building and only needs a simple pre-mapping of the desired locations.
The system is based on the same design principles as the Redpin[**?**] system that is to the author's best knowledge the first open source indoor localization system that used the intensity of the received signal (received signal strength – RSS). Unlike the Redpin system, MatMap uses a Naive Bayes Classifier whose performance has been shown to exceed the one of the Redpin system.

Keywords: indoor localization, mobile computing

1 Introduction

Precise location of a mobile device has become a crucial information for mobile devices application in recent years. While it is currently possible to obtain position estimates that are satisfactory for vast amount of applications in outdoor settings by using GPS, indoor localization still remains to be an open problem for many applications. For instance, if meter-level accuracy is desired the state of the art approach is to install additional hardware which then greatly improves the accuracy of such a system. However, installation of additional hardware is undesirable in many applications. In such setting the existing infractructure of WiFi access points needs to be utilized and the current state of the art methods provide room-level accuracy of position estimates.

In our work we focus on the design of a system that requires no additional hardware to be installed, is simple to implement and suitable to be used as part of a smartphone application while providing room-level accuracy.

1.1 Motivation

The campus of Faculty of Mathematics, Physics and Informatics of Comenius University in Bratislava can be a very confusing place, especially for a stranger.

Given the vast amount of rooms and halls that look all alike, one can get lost fairly easily (most of the time this is the case of freshman students). Our goal is to provide a framework for "official applications" of the Faculty which would be capable of detecting user's position and then navigating them around the Faculty.

1.2 Related work

In the past few decades a huge body of research has been done in the area of indoor localization. The systems it describes can be categorized by numerous factors into classes such as continuous (position is expressed as a coordinate in some sort of 2D or 3D coordinate system)[**?**], discrete (where the locations are represented in form of distinct labels, such as rooms)[**?**], those that rely on existing infrastructure (such as WiFi access points)[**?**] and those that rely on specific hardware that needs to be installed as part of their setup procedure (these solutions, such as [**?**], are usually based on ultrasound).

The system proposed in this paper could be categorized as discrete, since locations are distinct labes of places and would also belong to the category of those systems that rely on existing infrastructure. It shares many principles with the Redpin system [**?**] but uses a localization method described in [**?**] that yields more accurate results in therms of localization precision.

2 Model

2.1 Naive Bayes

As mentioned above, the system described in this work uses a naive Bayes classifier. All possible locations need to be known beforehand. We denote them $L_1, L_2, ..., L_N$ where N is the number of locations. These locations are estimated by a set of features $F_1(L_i), F_2(L_i), ..., F_M(L_i)$. These features ought to depend on the location only [1]. Under this assumption the probability that certain variables $f_1, f_2, ..., f_M$ were observed can be denoted as

$$P(L_i) = \frac{P(f_1|L_i) * P(f_2|L_i) * ... * P(f_M|L_i) * P(L_i)}{P(f_1) * P(f_2) * ... * P(f_M)}$$

We then estimate the location where these variables were measured as the most probable location, that is

$$L_{prob} = \underset{L_i}{\operatorname{argmax}} \frac{P(f_1|L_i) * P(f_2|L_i) * ... * P(f_M|L_i) * P(L_i)}{P(f_1) * P(f_2) * ... * P(f_M)}$$

We can note that the dominator would be the same for all L_i, thus we can simplify the formula to

[1] Hence the name naive, since this assumption often times does not hold.

$$L_{prob} = \underset{L_i}{\operatorname{argmax}} P(f_1|L_i) * P(f_2|L_i) * ... * P(f_M|L_i) * P(L_i)$$

In our definition of the problem we assume that all locations are equally likely to be observed. That means that $P(L_i)$ does not contribute to the overall result and thus we can remove it from the equation.

$$L_{prob} = \underset{L_i}{\operatorname{argmax}} P(f_1|L_i) * P(f_2|L_i) * ... * P(f_M|L_i)$$

In many applications of naive Bayes classification it is customary to take the logarithm of the product of probabilities and then express it as the sum of logarithms of probabilities. This is done in order to simplify the computation since probabilities are often very small numbers and taking their product might be troublesome to implement[2]. As we describe below this is not necessary in our use case.

2.2 Naive probability estimation

In some works that use naive Bayesian classifier the strength of the WiFi access point signal at a given location is completely ignored and thus the features depend on existence of a given access point. This doesn't yield particularly good results, as described in [?]. We try to avoid this problem by taking the signal strength of an access point into consideration.

The features $f_1, ..., f_M$ are in our case access points that were observed at the current scan. They can be uniquely described as a tuple $(ssid, signal)$ which we denote as $(f_i^{ssid}, f_i^{signal})$. In the same way we can express the recorded features for locations as $(L_i^{ssid}, L_i^{signal})$. It is necessary to note that since the features were recorded more than once in all locations they are grouped by L^{ssid} and the final L^{signal} is produced by taking the mean of all values L^{signal} that belong to the same L^{ssid}. The probability $P(f_i|L_i)$ can be then expressed as

$$P(f_i|L_i) = 100 - |L_i^{signal} - f_i^{signal}|$$

where $L_i^{ssid} = f_i^{ssid}$ and at the same time L_i^{signal} and f_i^{signal} are real values ranging from 0 to 100.

References

1. Filonenko, Viacheslav, Charlie Cullen, and James D. Carswell. "Indoor positioning for smartphones using asynchronous ultrasound trilateration." ISPRS International Journal of Geo-Information 2.3 (2013): 598-620. http://www.mdpi.com/2220-9964/2/3/598/htm

[2] Often times the biggest problems are with precision of floating point values, such as under-runs.

2. Bolliger, Philipp. "Redpin-adaptive, zero-configuration indoor localization through user collaboration." Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments. ACM, 2008. https://vs.inf.ethz.ch/publ/papers/bolligph-redpin2008.pdf
3. Kessel, Moritz, and Martin Werner. "SMARTPOS: Accurate and precise indoor positioning on mobile phones." MOBILITY 2011, The First International Conference on Mobile Services, Resources, and Users. 2011. http://www.thinkmind.org/download.php?articleid=mobility_2011_8_10_40203
4. Lin, Hsiuping, et al. "Enhanced indoor locationing in a congested Wi-Fi environment." (2009). http://repository.cmu.edu/cgi/viewcontent.cgi?article=1006&context=silicon_valley
5. Luo, Yan, Orland Hoeber, and Yuanzhu Chen. "Enhancing Wi-Fi fingerprinting for indoor positioning using human-centric collaborative feedback." Human-centric Computing and Information Sciences 3.1 (2013): 1-23. http://www.hcis-journal.com/content/3/1/2
6. Kovalev, Maxim. "Indoor Positioning of Mobile Devices by Combined Wi-Fi and GPS Signals." http://www.ipin2014.org/wp/pdf/5A-1.pdf