Mobile Habits: Inferring and predicting user activities with a location-aware smartphone

Andrei Papliatseyeu[1] and Oscar Mayora[2]

[1] University of Trento papliats@disi.unitn.it
[2] Create-Net oscar.mayora@create-net.org

Summary. In this paper we present a work in progress dedicated to the recognition and prediction of mobile user activities. In contrast with related projects that generally use GPS for localization, we employ a fusion of wireless positioning methods available in current smartphones (GPS, GSM, Wi-Fi). Our positioning system offers high availability and accuracy without dedicated calibration. We demonstrate how such a positioning information can improve place extraction algorithms and enable the recognition of the new types of user activities both indoors and outdoors. Besides that, the project addresses a number of open challenges in activity and place prediction, such as detection of behaviour changes, prediction of unseen places.

Key words: activity recognition, behaviour prediction, location awareness

1 Introduction

The knowledge of user activities and habits is a crucial factor for the development of highly personalized applications, that can be beneficial in many areas of daily life. First of all, the activity recognition is important in assisted living and health care systems. These methods can be used to support memory and planning capabilities of elderly [1], enable early detection of possible health problems [2, 3]. Mobile content providers can adapt the information delivery accordingly to the user preferences and the current context. There also are more general applications, like adaptation of mobile phone interface accordingly to the user's current activity [4]. Moreover, the prediction of user activity (and its location) can be used to improve routing and overall performance of wireless networks [5–7].

This project will explore and develop novel methods for recognition and prediction of user activities. Mobile smart devices present an ideal platform for this task; they usually possess considerable computational resources, a reach set of wireless communication and multimedia features. Assuming that people tend to always carry their phones along, we employ a modern smartphone as the main sensory and processing unit for learning user's behaviour.

A large number of research works have been dedicated to smartphone-based context recognition [8–10]. The authors focused primarily on detection of user's activity, paying little attention to the location; in most cases, only a single positioning technology was used, e.g. GPS [11, 12] or Wi-Fi [13]. However, any individual localization method has its limitations (GPS works only outdoors, Wi-Fi is rarely available outside of big cities) that impose certain constraints on the obtained results. Context prediction, in turn, enables the development of proactive features, such as taking preventive measures in health care, effective roaming in networks [5, 7], in-time traffic notifications [14], and others [15, 16]. The area of context prediction is not mature yet and offers many open challenges [15, 17].

In this project we propose to use a fusion of multiple positioning techniques available in modern smartphones (GPS, GSM, Wi-Fi, Bluetooth) for ubiquitous coverage and high accuracy of localization. We expect that such a positioning system will enable detection of new types of user activities and improve overall recognition accuracy. Another part of the project addresses the open issues of activity prediction, such as detection of changes in habitual behaviour, recognition of activities with complex periodicity and data uncertainty handling.

The paper proceeds as follows. Section 2 describes related work and some of the current challenges. Section 3 presents our objectives and approach. We conclude with a brief description of the current state and future work on the project.

2 Related work

2.1 Positioning frameworks

There is no ideal positioning technology. GPS provides an accuracy better than 10m, but only outdoors and with non-obstructed view of sky. Median positioning error of a state-of-the-art Wi-Fi positioning system can be less than 2m, but Wi-Fi coverage is very limited in less populated areas and developing countries. GSM-based solutions provide high coverage but have low accuracy (hundreds of meters). Bluetooth can be used for sub-room-level positioning, but stationary Bluetooth-enabled devices are not widely spread yet.

The current approach to this problem is to combine the data from multiple positioning sensors in order to increase the overall coverage of the system and improve its accuracy. This process is called *sensor fusion* and defined as

> ...the use of multiple technologies or location systems simultaneously to form hierarchical and overlapping levels of sensing... [It] can provide aggregate properties unavailable when using location systems individually. [18]

In the last decade, a number of location frameworks featuring sensor fusion methods have been proposed [19–22]. Some of these works were rather generic [19, 22], while others were targeted specifically for mobile devices with limited resources [20, 21]. PlaceLab [20] uses multiple wireless technologies for positioning (GPS, GSM Cell-ID, Wi-Fi). It is implemented in Java with some native code, and is available for different platforms. All the measurements are converted to a common physical coordinate system. Data fusion (called by the authors as "conflict solving") can be done using a simple Venn diagram-like method or more resource-demanding particle filtering approach. However, the training of PlaceLab is difficult, as it requires extensive mapping of the RSSI and beacon-ID data to the ground-truth position provided by GPS; this becomes a particularly problematic task indoors because GPS is not available there. Also, the evaluation has shown that the conflict solving method does not increase the positioning accuracy, but instead improves the coverage at the cost of lower accuracy [20, p. 129]. Another localization framework, Location Stack, developed by Hightower et al. [19] represents a layered OSI-like architecture of a positioning system. The data from multiple positioning sensors are first converted to an internal representation in Cartesian coordinates; Bayesian particle filtering and motion modelling are then used for probabilistic data fusion. Although we do not use Location Stack as a positioning system, our work is closely related to its top layers, namely, Activities and Intentions, that were not considered by the authors.

Depending on their type, positioning systems provide either *physical* or *symbolic* location [18]. The systems of the first type report the location as coordinates (absolute or relative), while the second type systems output a label or a name, associated with the place. Most of the current positioning systems, except GPS, are symbolic; in order to obtain physical coordinates from them one needs a database which maps symbolic locations to their coordinates. However, this approach requires considerable calibration efforts and is not well-scalable. The inverse conversion from physical coordinates to a meaningful place name (*place extraction*) is not a straightforward process; some of its aspects remain a challenge.

2.2 Place extraction and prediction

Kang et al. define *place* as

> ...a locale that is important to an individual user and carries important semantic meanings such as being a place where one works, live, plays, meets socially with others, etc.[23]

A number of algorithms have been developed for recognition of personally important (or frequent) places. Ashbrook and Starner [11] made use of poor indoors GPS reception: the loss of GPS signal for more than 10 minutes was treated as being in an important place; hierarchical clustering was then used to identify subspaces. The method inherits the drawbacks of GPS, namely, limited availability in urban environment, and is not suitable for place recognition indoors. Laasonen et al. [24] offered a method for place extraction basing on GSM cell transitions; the accuracy was limited due to low resolution of GSM Cell-ID positioning. An adaptive clustering approach proposed in [23] used time and distance thresholds for computationally effective place extraction; the method was tested on Wi-Fi positioning traces obtained via PlaceLab. Similarly, Rekimoto et al. [13] used a custom Wi-Fi keychain logger and offline k-means clustering. A different view of place extraction was presented by Hightower et al. [25]. In order to avoid the limitations of single-sensor positioning and the calibration requirement of PlaceLab, they offered a BeaconPrint clustering algorithm that simultaneously uses Wi-Fi and GSM fingerprints for place recognition.

To our best knowledge, all current place extraction methods apply a binary decision rule to classify a place either as important or not. We argue that such an approach considerably reduces the flexibility of personalized location-aware applications. Instead, we propose to use a more gradual classification, which preserves the “importance” rank of the place (see Section 3.1).

Prediction of the user’s future location is often an integral part of the place recognition works. Ashbrook and Starner [11] used a Markov model to predict user movements between places. However, the model was updated offline and the time of the move was not considered. A similar work, but using Bayes predictor, was conducted by Krumm and Horvitz [26]. Their approach addresses the problem of prediction of “never-seen-before” locations, but is not applicable for a mobile device due to the high computational complexity. Song et al. [27] have compared four prediction algorithms in a Wi-Fi network, and shown that a simple Markov model performed almost as good as other, more complicated methods. Mayrhofer in his PhD thesis [16] addressed a generalized problem of context prediction on a mobile device. The consideration of only partial history (using sliding window) and prediction of abstract context instead of its particular features can be seen as limitations of the work.

The places that have never been visited before represent significant difficulties for prediction. This problem has been addressed by open-world modelling approach [26] and by shortening the context until the algorithm is able to provide a prediction [27]. Another challenge is to relax the *stability assumption* [27], which implies that the user behaviour does not change through the time. For example, a college student’s behaviour alters rapidly at the beginning of a new semester as the class schedule changes, and the predictor can take the whole new semester to adapt [11]. A typical approach is to assign recent events more weight (however, with little effect [27]). In this paper we present novel methods addressing both of these challenges.

2.3 Activity recognition

A number of projects have addressed the problem of inferring user’s activity from position. A two-step hidden Markov model (HMM) and the variance of Wi-Fi signal strength were used in LOCADIO project [28] to detect whether the user is in movement. Patterson et al. [12] were able to recognize the transportation mode of the user (by bus, by car or by feet) using an online unsupervised Bayesian classifier. A similar task has been done in [29] for GSM network; they have demonstrated that unsupervised HMM and a supervised ANN have similar recognition accuracy (about 80%). Sohn et al. [30], in turn, achieved 85% accuracy by applying a boosted logistic regression with one-node decision tree. A large-scale study of user activity patterns has been performed by Reality Mining project [8]. One hundred users with ContextPhone-running smartphones [10] were building the biggest known publicly available DB of contextual data, over the period of 9 month. The project enabled the researchers to analyze many different perspectives of stand-alone and cooperative user behaviour, activity patterns, etc. Although one of the key ideas of the project was to use both Cell-ID and Bluetooth proximity to other devices, so that the two techniques could "augment each other for location and activity inference", the user location data were very coarse-grained due to the limitations of the ContextPhone platform [10].

3 Our objectives and approach

The aim of the project is the development of methods for recognition and prediction of mobile user activities by means of sensors and resources available in a modern smartphone. In particular, we focus on utilization of wireless positioning techniques and user movement history obtained from their fusion. The project objectives are the following.

- Analyse and, if necessary, develop place extraction methods for multi-sensor positioning system, using the resources available in modern smartphones.Address the issues of localization uncertainty, privacy and minimization of user distraction.
- Analyse the types of activities that can possibly be inferred from user's current location and movement history; develop corresponding classification methods, considering the probabilistic nature of location estimates and their varying accuracy. Identify auxiliary sources that can improve classification accuracy (e.g. time, day of the week, recent calls, etc).
- Analyse the applicability of machine learning techniques for recognition of periodic patterns in the user behaviour and prediction of future user activities (intentions).
- Evaluate the performance of the developed methods via user study.

3.1 Hybrid positioning and place recognition

As it has been shown in the Section 2.3, almost all of activity recognition projects employ a single positioning technology, usually GPS. The novelty of our approach is in the utilization of a fusion of positioning methods that

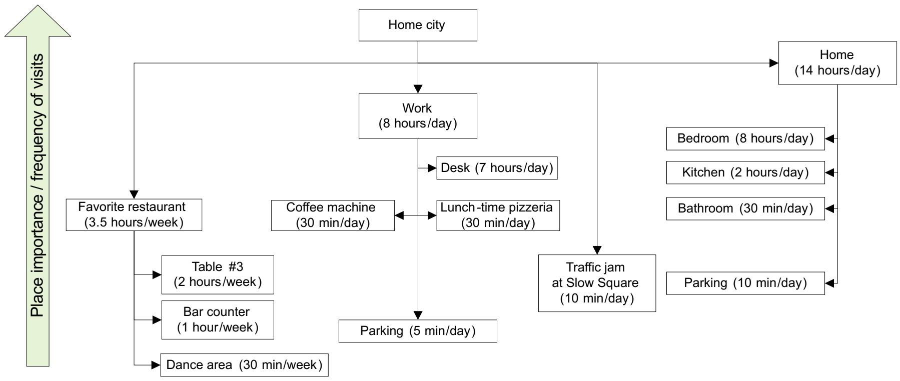

Fig. 1. An example of a user's weekly movement pattern taking into account places "importance".

provides higher coverage and accuracy. These factors are expected to enable recognition of new types of activities and improve recognition performance. In order to avoid expensive and non-scalable calibration, we adapt a hybrid positioning approach: places are recognized by their raw fingerprints; geographic coordinates are assigned only when available. Our approach is similar to that of Hightower et al. [25]; however, while they have used purely symbolic positioning, we augment the fingerprints with GPS coordinates when possible. This enables the use of coordinates-based inference, like reverse geocoding, obtaining a list of nearby businesses, etc.

The increased availability of positioning information enables us to recognize personally important places both outdoors and indoors with high accuracy. Consideration of places at different scale is a challenging task, because inter-place distances indoors and outdoors vary by orders of magnitude, approaching positioning error when indoors [31]. We propose to augment the recognized places with the "importance" rank. Although most of the current place extraction algorithms already use some kind of ranking in their intermediate stages, the final step is usually a threshold-based binary decision on whether a place is significant or not. We believe that consideration of "Work" place being more important than "Post box" might be more beneficial than the traditional assumption they are equally significant (see Figure 3.1). This approach alleviates the choice of proper threshold value, because in our approach the latter only defines the lower bound of places to be considered.

3.2 Activity recognition

An important issue for a symbolic positioning system we employ is the labelling of discovered places with personally significant names. For example, the user would probably prefer to label his/her house as "Home" instead of address or geographic coordinates. One possible approach can be to request the user to input place's label manually. However, this is distracting and can easily become irritating. Another approach for automatic place labelling can be beneficial for the users who schedule their meetings using phone's agenda software. Meeting description usually includes some kind of "location" field, where the user enters personalized name of the meeting place. Observing the user's place at the meeting times for a certain period, the labelling algorithm can assign the corresponding name to the place, along with some degree of confidence that this name is appropriate. Unfortunately, sometimes it is impossible to infer the name of the place without direct user input. For such cases we propose to separate place labels into two categories: private and public. The latter can be uploaded to the Internet and made available for other users, so that they can avoid repetitive manual labelling of public places.

Accurate positioning information augmented by basic inference rules can substantially widen the range of recognizable activities. For example, given GPS coordinates of a place, the system may apply reverse geocoding followed by white pages lookup and retrieve the name of the business located at this place and its category (e.g., "pub") [32]. Then, relying on timing and the observations of internal movements at this place, the system can infer additional information about the user. In the "pub" example, being there at the same time period longer than 4-6 hours, more than 4-5 times a week for some weeks, means that the user is either a *barman* (if staying at the same place during the whole period) or a *waiter/waitress* (moving between a number of places, i.e. tables); in other cases the user can be considered as a *visitor*.

Hybrid positioning also enables the graceful degradation of system performance: if Wi-Fi is not available, the system is unable to detect internal movements but still knows the GPS coordinates of the place (and, possibly, its category); if the GPS signal is poor, the system can still recognise the place by its GSM and/or Wi-Fi fingerprints.

3.3 Place and activity prediction

Existing methods for the prediction of user movements, described in Section 2.2, usually model user behaviour as a graph, where nodes represent significant places and edges represent routes. However, this model misses such important factors as time and periodicity. Indeed, the probability of going from "Home" to "Work" during weekend is arguably lower than during working days. Moreover, some activities have long (e.g., visiting a dentist) or complex periods (e.g., visiting friends' house when anyone of them has birthday). For the recognition of periodic activities, we plan to further explore the Fourier transform approach demonstrated in the Reality Mining project [8].

The activity prediction methods presented in related research are also affected by the so-called *"schedule change" problem.* For example, the behaviour of a university student is to a large extent defined by the timetable of classes [11]. During the semester the system learns the schedule and reaches reasonable prediction accuracy. However, when a new semester begins and the schedule changes, the system might take the whole semester to retrain. Typical workarounds are either to shorten the training history or to weight it, assigning more importance to the recent data [27]. Unfortunately, both of these approaches reduce the system capability of prediction of rare events, and should be used only when the need for retraining is detected. We propose a self-evaluation approach for detection of behaviour change. For an "always on" system it is possible to constantly check whether its forecasts come true within certain time interval. If the number of errors increases over its usual value and remains high for some period, it might be an indication of a schedule change. In this case the system should facilitate quicker retraining by application of the weighting method described above.

Availability of positioning history enhanced with categorical information (see Section 3.2) can offer a step towards the prediction of previously unseen places. Let us consider a user whose activity profile exhibits a tendency to spend lunchtime in places categorized as "pizzeria". In this case, even if the user is in a different city, the system may utilize GPS coordinates and web search to get a list of nearby pizzerias and treat them as possible location at lunchtime.

3.4 Project plan and current state

Currently, we have finished the literature review and are implementing the data collection application based on POLS framework [33] with a GSM smartphone and an external GPS receiver. Ground truth is to be provided by the user via graphical interface. The project will proceed with a data collection phase: 9 month by the author and, after ensuring the robustness of the logging system, 2 weeks by lab members. For the place extraction, we will adapt the BeaconPrint algorithm [25] augmented with the place importance measure described in Section 3.1.

We have planned a number of experiments for testing our hypotheses and performance of learning methods. First of all, the positioning traces will be used to compare stand-alone and joint availability of different positioning technologies. Then, we plan to analyse the performance of activity recognition methods (including k-means clustering, k-nearest neighbour, naïve Bayes, hidden Markov models and simple neural networks). For the detection of the behavioural patterns, we will evaluate clustering algorithms, fast Fourier transform and self-organizing Kohonen maps. Finally, the prediction methods that take into account the time will be additionally evaluated with a separate dataset collected on city bus routes.

4 Conclusion

This paper presents our approach to recognition and prediction of user's activities using a location-aware smartphone. We propose a multi-sensor localization system which offers almost ubiquitous coverage without prior calibration.

The system uses hybrid position estimates based on GSM and Wi-Fi fingerprinting, optionally augmented with GPS coordinates when they are available. We argue that such a system can increase the range of detectable activities. We also demonstrate how basic common-sense reasoning can further improve recognition. Besides that, the article describes our approach to the recognition and automatic labelling of personally important places, prediction of periodic activities and behaviour change detection.

References

[1] C.D. Kidd, R.Orr, G.D. Abowd, C.G. Atkeson, I.A. Essa, B. MacIntyre, E. Mynatt, T.E. Starner, and W. Newstetter. The Aware Home: A Living Laboratory for Ubiquitous Computing Research. In *Proc. CoBuild 99*, volume 1670 of *LNCS*, pages 191–198. Springer, 1999.

[2] D. Tacconi, O. Mayora, P. Lukowicz, B. Arnrich, C. Setz, G. Troester, and C. Haring. Activity and Emotion Recognition in the P-cube Framework to Support Early Diagnosis of Psychiatric Diseases. In *Proc. 2nd Int'l Conf. on Pervasive Computing Technologies for Healthcare*, 2008. (To appear).

[3] H. Kautz, L. Arnstein, G. Borriello, O. Etzioni, and D. Fox. An overview of the assisted cognition project. In *Proc. AAAI-02 Worskhop on Automation as Caregiver: The Role of Intelligent Technology in Elder Care*, pages 60–65, 2002.

[4] D.S. Weld, C. Anderson, P. Domingos, O. Etzioni, K. Gajos, T. Lau, and S. Wolfman. Automatically personalizing user interfaces. In *Proc. IJCAI-03*, 2003.

[5] G. Liu and G. Maguire. A class of mobile motion prediction algorithms for wireless mobile computing and communications. *Mobile Networks and Applications*, 1(2):113–121, 1996.

[6] J.-K. Lee and J.C. Hou. Modeling steady-state and transient behaviors of user mobility: formulation, analysis, and application. In *Proc. 7th ACM Int'l Symposium on Mobile ad hoc networking and computing*, pages 85–96. ACM, 2006.

[7] W. Su, S.J. Lee, and M. Gerla. Mobility prediction and routing in ad hoc wireless networks. *Int. J. Network Management*, 11(1):3–30, 2001. ISSN 1099-1190.

[8] N. Eagle and A. Pentland. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):255–268, 2006.

[9] I. Mulder, H. ter Hofte, R. Otte, and P. Ebben. Collecting user experience in context with Xensor for Smartphone. In *Proc. Int'l Symposium on Intelligent Environments*, 2006.

[10] M. Raento, A. Oulasvirta, R. Petit, and H. Toivonen. ContextPhone: A Prototyping Platform for Context-Aware Mobile Applications. *IEEE Pervasive Computing*, 4(2):51–59, 2005.

[11] D. Ashbrook and T. Starner. Using GPS to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous Computing*, 7(5):275–286, 2003.

[12] D. Patterson, L. Liao, D. Fox, and H. Kautz. Inferring high-level behavior from low-level sensors. In Anind K. Dey, Albrecht Schmidt, and Joseph F. McCarthy, editors, *Proc. UbiComp 2003*, volume 2864 of *LNCS*, pages 73–89. Springer, 2003.

[13] J. Rekimoto, T. Miyaki, and T. Ishizawa. LifeTag: WiFi-Based Continuous Location Logging for Life Pattern Analysis. In *Proc. LoCA'07*, volume 4718 of *LNCS*, pages 35–49. Springer, 2007.

[14] K. Zhang, K. Torkkola, H. Li, C. Schreiner, H. Zhang, M. Gardner, and Z. Zhao. A Context Aware Automatic Traffic Notification System for Cell Phones. In *Proc. 27th Int'l Conf. on Distributed Computing Systems Workshops*. IEEE, 2007.

[15] E. Mynatt and J. Tullio. Inferring calendar event attendance. In *Proc. IUI'01*, pages 121–128, 2001.

[16] R. Mayrhofer. *An Architecture for Context Prediction*. PhD thesis, Johannes Kepler University of Linz, Austria, 2004.

[17] R. Mayrhofer. Context Prediction based on Context Histories: Expected Benefits, Issues and Current State-of-the-Art. In *Proc. ECHISE'05*, 2005.

[18] J. Hightower and G. Borriello. Location systems for ubiquitous computing. *Computer*, 34(8):57–66, 2001.

[19] J. Hightower, B. Brumitt, and G. Borriello. The location stack: a layered model for location in ubiquitous computing. *Proc. 4th IEEE Workshop on Mobile Computing Systems and Applications*, pages 22–28, 2002.

[20] A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, J Tabert, P. Poweldge, G Borriello, and B Schilit. Place Lab: Device Positioning Using Radio Beacons in the Wild. In *Proc. Pervasive'05*, volume 3468 of *LNCS*, pages 116–133. Springer, 2005.

[21] J. Nord, K. Synnes, and P. Parnes. An architecture for location aware applications. In *Proc. HICSS'02*, pages 3805–3810, 2002.

[22] F. Kargl and A. Bernauer. The COMPASS Location System. In *Proc. LoCA 2005*, volume 3479 of *LNCS*, pages 105–112. Springer, 2005.

[23] J.H. Kang, W. Welbourne, B. Stewart, and G. Borriello. Extracting places from traces of locations. *ACM SIGMOBILE Mobile Computing and Communications Review*, 9(3):58–68, 2005.

[24] K. Laasonen, M. Raento, and H. Toivonen. Adaptive On-Device Location Recognition. In *Proc. Pervasive'04*, volume 4, pages 287–304. Springer, 2004.

[25] J. Hightower, S. Consolvo, A. LaMarca, I. Smith, and J. Hughes. Learning and Recognizing the Places We Go. In *Proc. UbiComp'05*, volume 3660 of *LNCS*, pages 159–176. Springer, 2005.

[26] J. Krumm and E. Horvitz. Predestination: Inferring Destinations from Partial Trajectories. In *Proc. UbiComp'06)*, volume 4206 of *LNCS*, pages 243–260. Springer, 2006.

[27] L. Song, D. Kotz, R. Jain, and X. He. Evaluating next-cell predictors with extensive Wi-Fi mobility data. *IEEE Transactions on Mobile Computing*, 5 (12):1633–1649, 2006.

[28] J. Krumm and E. Horvitz. LOCADIO: Inferring Motion and Location from Wi-Fi Signal Strengths. In *Proc. MobiQuitous'04*, pages 4–13. IEEE, 2004.

[29] I. Anderson and H. Muller. Practical Activity Recognition using GSM Data. Technical Report CSTR-06-016, Department of Computer Science, University of Bristol, 2006.

[30] T. Sohn, A. Varshavsky, A. LaMarca, M.Y. Chen, T. Choudhury, I. Smith, S. Consolvo, J. Hightower, W.G. Griswold, and E. de Lara. Mobility Detection Using Everyday GSM Traces. In *Proc. UbiComp'06*, volume 4206 of *LNCS*, pages 212–224. Springer, 2006.

[31] R. Aipperspach, T. Rattenbury, A. Woodruff, and J. Canny. A Quantitative Method for Revealing and Comparing Places in the Home. In *Proc. UbiComp 2006*, volume 4206 of *LNCS*, pages 1–18. Springer, 2006.

[32] G. Castelli, M. Mamei, and A. Rosi. The Whereabouts Diary. In *Proc. LoCA'07*, volume 4718 of *LNCS*, pages 175–192. Springer, 2007.

[33] Privacy-Observant Location System. Online: `http://pols.sourceforge.net/`. (Accessed 2008.05.07).