Situated Public Displays that learn the structure of a building

Hans Jörg Müller, Oliver Paczkowski and Antonio Krüger

As display prices fall, public displays rapidly become ubiquitous within organizations. Many organizations already have displays installed in their entrances. Finally, from every place within a building a public display will be visible. However, how they could be used best is still unclear. Today most displays show static Powerpoint presentations or video loops. These presentations are neither adapted to the spatial and temporal context of the display nor to the users watching them.

Figure 1: One Display installed in the main entrance.

We suggest to use these displays as information systems that show discrete information as well as continuous information about and from within an organization. As people pass by these displays, they have a quick glimpse to get the newest information about the organization. To provide the information that is most interesting for the people in front of the display, the public displays need to become situated, i.e. become aware of their spatial, temporal and social context. Since we consider it infeasible that a model of context is provided manually, our system will automatically configure itself to its context and will not require any manual configuration except for mounting the displays on the wall.

In this paper we focus on modeling the spatial context of situated public displays. We present an approach to learn the relative location of the displays and important places within a building by collecting interaction data from the occupants.
We are currently building a prototype of such a situated public display system within our university department. The information displayed consists of both discrete information chunks that are authored by people in the department and continuous information streams that come from sensors, databases or web pages. Information chunks are mostly events (like talks), job offers (like open diploma theses) and current status information about the organization (like new members). Information streams are open facilities, the cafeteria menu, the weather, and the next bus departures.
We currently have three displays installed within our department and one in another. These displays are located at entrances and main hallways, where many people pass by. According to our experiences, most people have a quick glimpse at the displays quite frequently. We also work on interaction with situated public displays, for example interaction with mobile phones.

Users can install a Java MIDlet on their mobile phone to connect to the displays via Bluetooth and get detailed information on topics displayed. Those users can agree that every time they pass a display, a timestamp is logged, and also provide the social group they belong to (like geoinformatics student) so the displays can adapt to them.

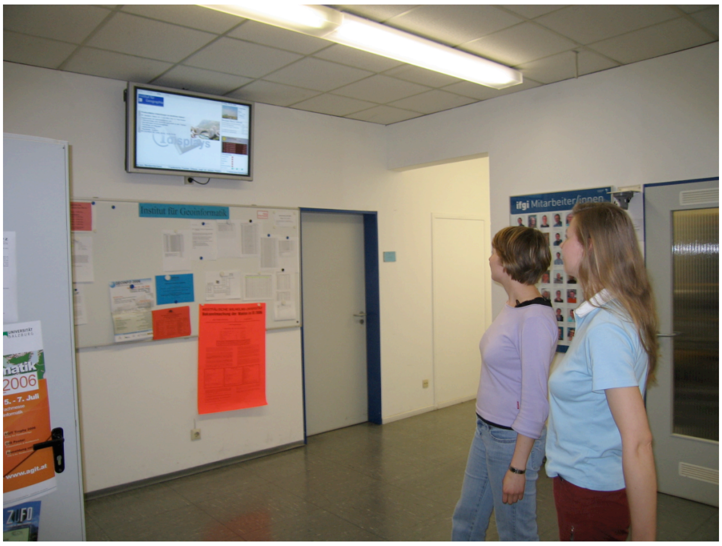

Figure 2: One Display installed in a main hallway.

Our system learns a topological model of places within the building from the data collected via Bluetooth. Within the model the displays are represented as nodes. If it is possible to walk from one display to another without passing a third display, these displays are connected with edges. The edges are labeled with the mean walking time for the direct way between the displays. In addition, important rooms like lecture halls and seminar rooms are automatically added to the edges representing the hallways that they are connected to.

The structure of the building is learned in two steps. In the first step, the relative position of displays along with the walking time needed to traverse between them is learned. In the second step, the topology of important places is learned.
To gather data needed for learning, each display continuously scans for nearby Bluetooth devices. We exploit the fact that mobile phones are personal devices, and one phone is usually only carried by the same person. In addition to this, most people always carry their mobile phone with them. Since a scan for Bluetooth devices takes about 11 seconds, this period determines the temporal resolution of our logging method. The range of our Bluetooth antennae is about 10 meters.
If during a scan a device that opted in for the service is found, the device id is stored in a cache memory along with the display id and a timestamp. At any point in time, only the display where the device has been seen last needs to be cached.

The first learning algorithm, that learns the relative position of displays along with the walking time needed to traverse between them, is executed every time a device is encountered at a display. From the cache, the display where the device has been seen last is retrieved. The time difference between the two readings along with the two display ids is then stored as a traversal time in a database. Thus, for each pair of displays a collection of anonymous traversal times is created. From these traversal times, only a small percentage originates from the direct way between the displays, because people can walk alternative ways, or even pass one display, go home, and pass the other display the next week. In addition, even for people that take the direct path between two displays the traversal times can vary. To deal with this problem we use the Expectation-Maximization (EM) algorithm to cluster the traversal times and identify multiple paths. We assume that for each possible path the traversal times are distributed according to the normal distribution. We propose to use three normal distributions for each pair of displays, so the direct way, one indirect way and a collection of all other ways can be identified. The EM algorithm is an unsupervised clustering algorithm. It adjusts the mean and standard derivation of these distributions such that the probability that these distributions have generated the data is maximized. More concretely, in a first step (E-Step), the expected value that each of the traversal times is generated by a certain path is calculated. This step uses the current hypothesis for the mean and standard derivation of all paths. In a second step (M-Step), a new maximum likelihood hypothesis for the means and standard derivations of all distributions is generated. This step assumes that each instance was generated by the distribution identified in the first step. These two steps are repeated until a local optimum is reached.

The second learning algorithm learns the topology of important places within the building. We use the same cache that is used for the first one, but assume that the social groups that certain people belong to are known. Since event information along with the date, place and social group that is probably interested in this event is available on the displays, the topological structure of these places can be learned from the people that go there to attend events. To achieve this, at the beginning of each event the cache is searched for the display ids where each member of the corresponding group has been seen last. From these, those members who are walking around during the event and caused Bluetooth readings on different displays are removed. The other members of the group are assumed to be attending the event, and the location of the place is assumed to be between the displays where those people have been seen last. In addition, when members of the group are continuously within the scan range of a display during an event, the corresponding place is assumed to be in immediate vicinity of the display. With this algorithm we make no assumptions on distances between rooms and the displays but solely on the topological relationships among them.

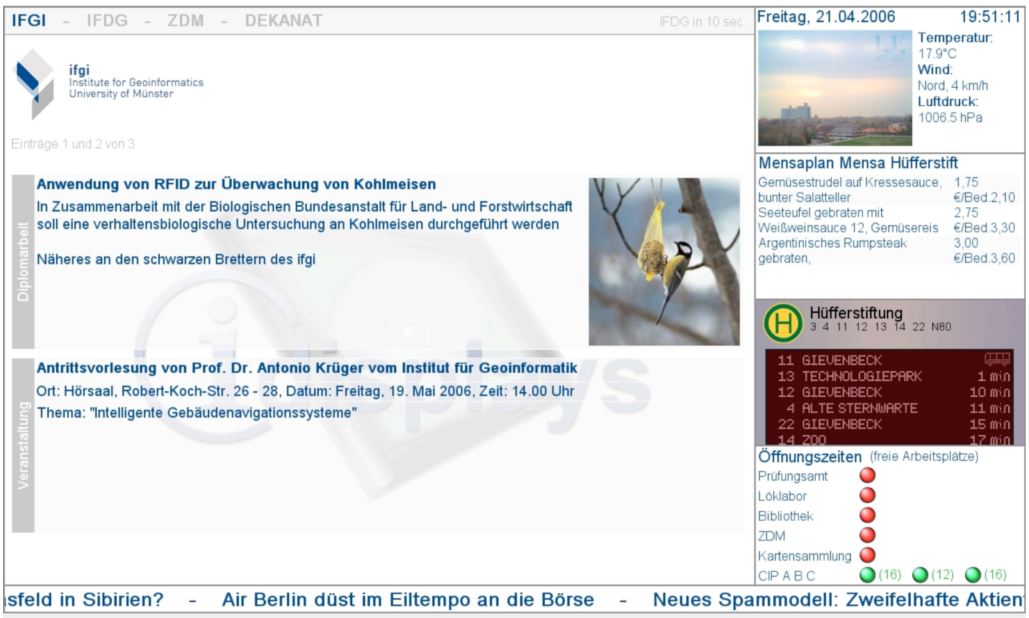

Figure 3: Display content with information chunks regarding the organization on the left, information streams regarding the immediate surroundings on the right and a ticker with general information on the bottom.

To include direction data into the topological model, additional sensors available at the displays, like cameras watching people in front of the displays, could be used. By combining optical flow with our Bluetooth information, for example, the displays could learn about the relative direction of other displays and places.

Using our model, it is possible to adapt information presentation on the displays to the spatial context. Thus, information regarding a talk or meeting could only be displayed on the displays next to the location of the event and not on displays at completely different locations. In addition to this the spatial model of the building will serve as the basis of more elaborate models describing the spatio-temporal and social behavior of the occupants of the building. For example, it is possible to predict both the trajectories that people will take through the building and their goal. Using trajectory prediction, one could then plan entire presentations for people traversing the building. On the first display they pass only headlines could be presented, on the second display more detailed information could be shown and on the third display an interaction possibility could be available. Using goal recognition, information regarding these goals, like the abstract of the lecture one is going to attend, could be shown.