UNI KASSEL VERSITÄT

ENDOWED CHAIR OF THE HERTIE FOUNDATION
Knowledge and Data Engineering
ELECTRICAL ENGINEERING & COMPUTER SCIENCE, UNIVERSITY OF KASSEL

Resource-Aware On-Line RFID Localization Using Proximity Data

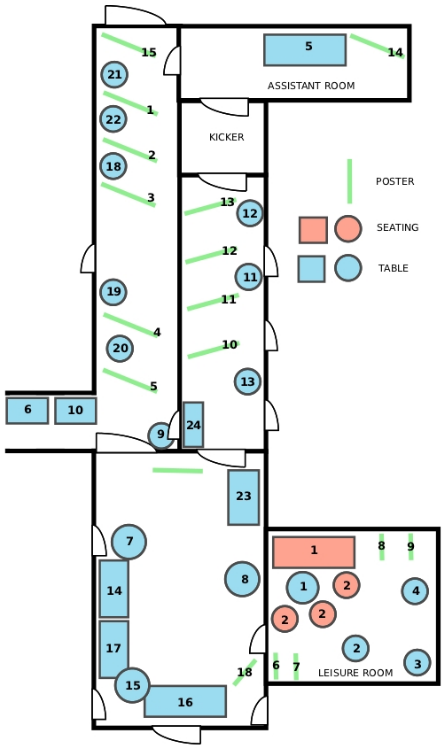
Christoph Scholz[1], Stephan Doerfel[1], Martin Atzmueller[1], Andreas Hotho[2], Gerd Stumme[1]

[1]Knowledge and Data Engineering Group
University of Kassel, Germany

[2]Data Mining and Information Retrieval Group
University of Würzburg, Germany

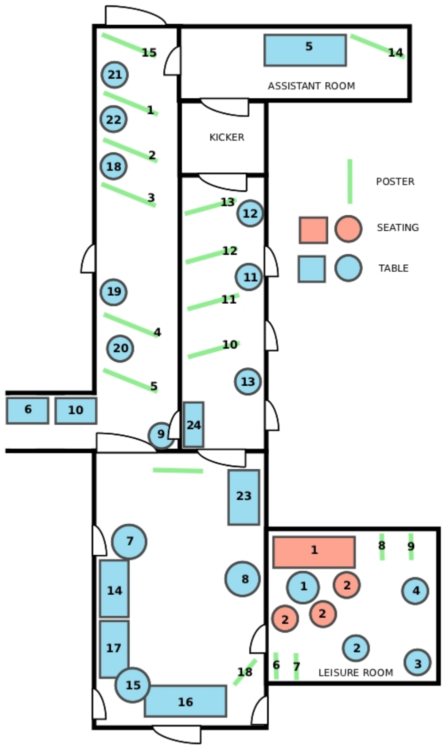
Agenda

1. Application Context
2. Method
    - Benchmark
    - Social Boosting
3. Dataset
4. Results
5. Summary

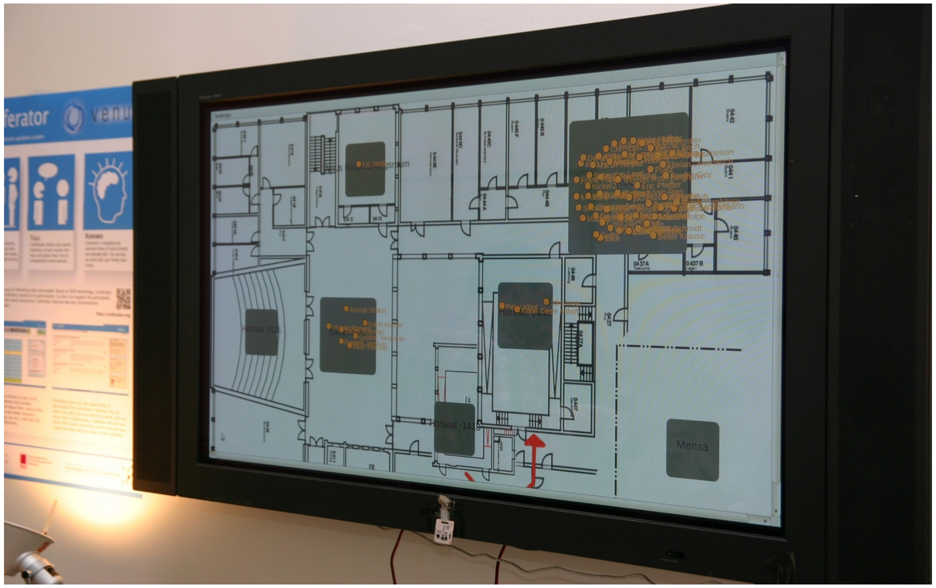


Agenda

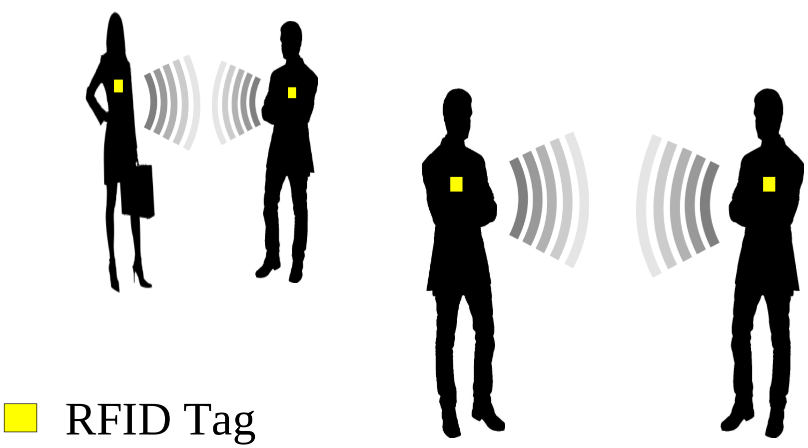
1. **Application Context**
2. Method
   - Benchmark
   - Social Boosting
3. Dataset
4. Results
5. Summary



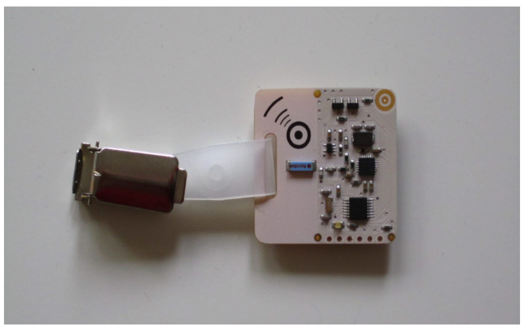
Localization of Conferences



- development of a social network for conferences (www.conferator.org)
- offers contact history for participants
- analysis of contact data
- indoor-localization (room level)



Proximity Tag



provided by

SocioPatterns

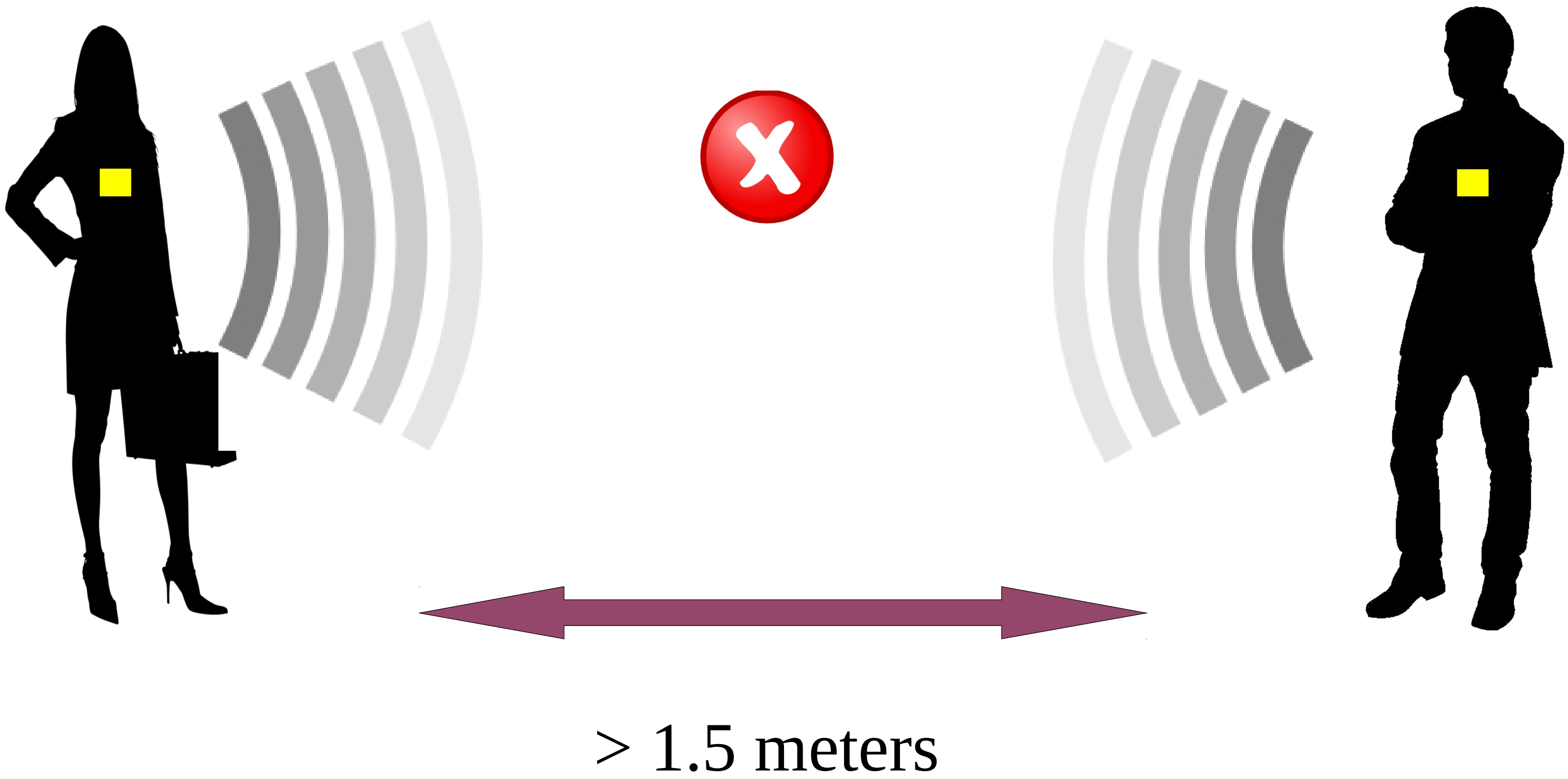
- active RFID tag
- detects other proximity tags within a range of up to 1.5 meters
- done by sending out special proximity packages
- sends one tracking package in four different signal strengths every two seconds
- cannot store information

provided by

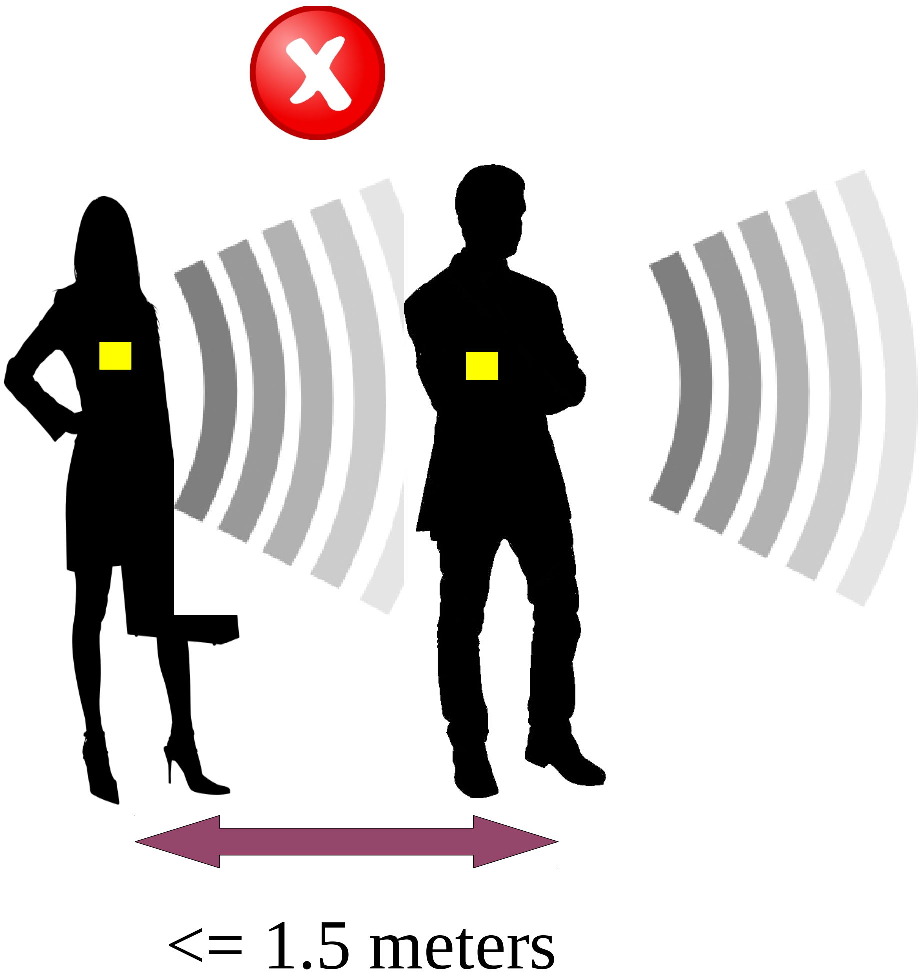
bitmanufaktur

- receives RFID signals from proximity tags and forwards them to a central server
- does not provide information like Angle of Arrival (AoA) or Received Signal Strength (RSS)

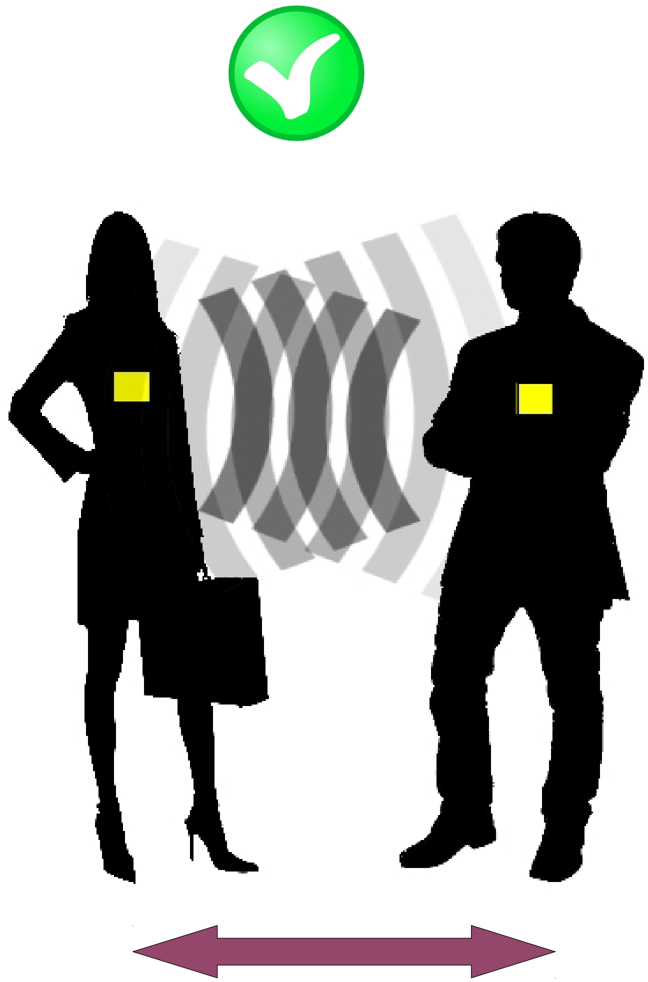
Contact Recognition



Contact Recognition



Contact Recognition



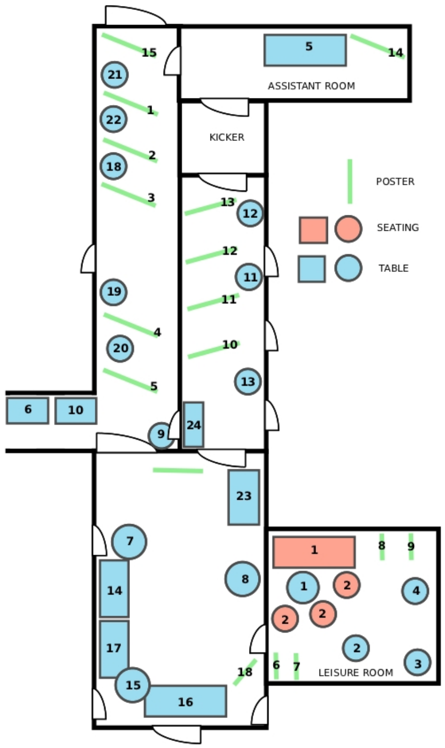
<= 1.5 meters

Contribution

- first explicit analysis that contact information helps to improve localization accuracy
- use of real-time data
- evaluation of state-of-the-art machine learning techniques for room-level localization
- resource-aware application
  - use of limited number of readers
  - use of cost-effective technology

Agenda

1. Application Context
2. **Method**
   - **Benchmark**
   - Social Boosting
3. Dataset
4. Results
5. Summary



Localization Using Characteristic Vectors

Assumption

The number of packages an RFID reader receives is significantly dependent on the position of the sending RFID tag.

↓

We determine a set of characteristic vectors for each room in the conference area.

Characteristic Vector

Reader 1 | Reader 2 | Reader 3 | Reader 4

time window l

$$charVec(1135,10)=(0,0,1,3,\quad 2,3,4,5,\quad 0,0,0,0,\quad 2,3,3,3)$$

1 2 3 4 | 1 2 3 4 | 1 2 3 4 | 1 2 3 4

RFID tag id

signal strength s

3 tracking packages were received by reader 1 from proximity tag 1135 with signal strength 4 within the last 10 seconds.

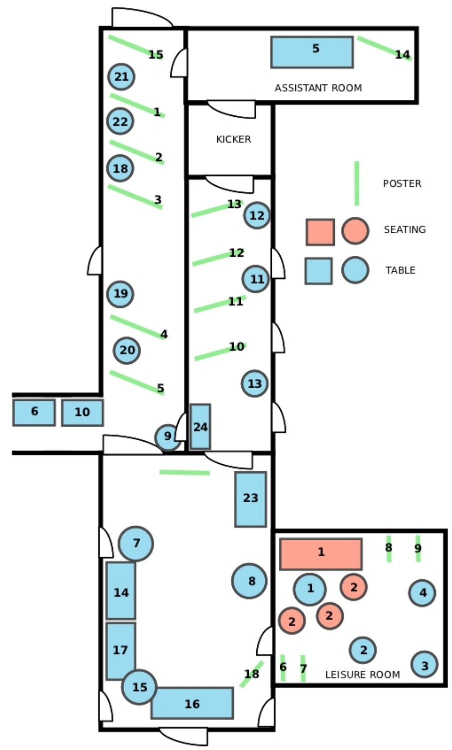
Benchmark

four state of the art machine learning methods:

- Naive Bayes (NBAY)
- K-Nearest Neighbor (KNN)
- Support Vector Machines (SVM) (radial kernel)
- Random Forest (RF)

Agenda

1. Application Context
2. **Method**
   - State of the Art
   - **Social Boosting**
3. Dataset
4. Results
5. Summary



Mean-Approach

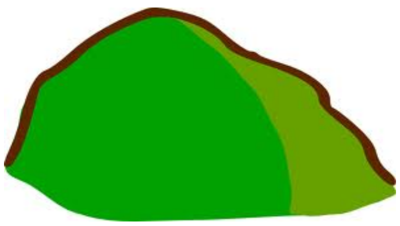
Goal: Prediction of user *u*'s position at time *t*,
given time window l

1. determine all contacts *C(u)* of user *u* at time *t*
2. set

$$charVec_{mean}(u,l)=\frac{charcVec(u,l)+\sum_{i\in C(u)} charVec(i,l)}{1+|C(u)|}$$

3. use $charVec_{mean}(u,l)$ as input vector for the prediction algorithm

Max-Approach



Goal: Prediction of user *u*'s position at time *t*, given time window l

1. determine all contacts *C(u)* of user *u* at time *t*
2. set

$$charVec_{max}(u,l)=\left(max_{i\in C(u)\cup\{u\}}\left\{charVec(u,l)_1\right\},\ldots,max_{i\in C(u)\cup\{u\}}\left\{charVec(u,l)_{4r}\right\}\right)$$

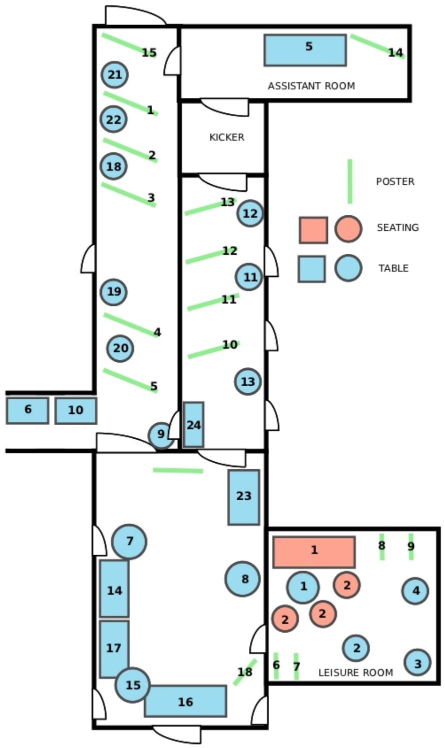
3. use $charVec_{max}(u,l)$ as input vector for the prediction algorithm

Vote-Approach

1. Determine the position (room) for user *u* and all his contacts with the prediction algorithm.
2. User *u*'s position is the majority vote among all these predictions.

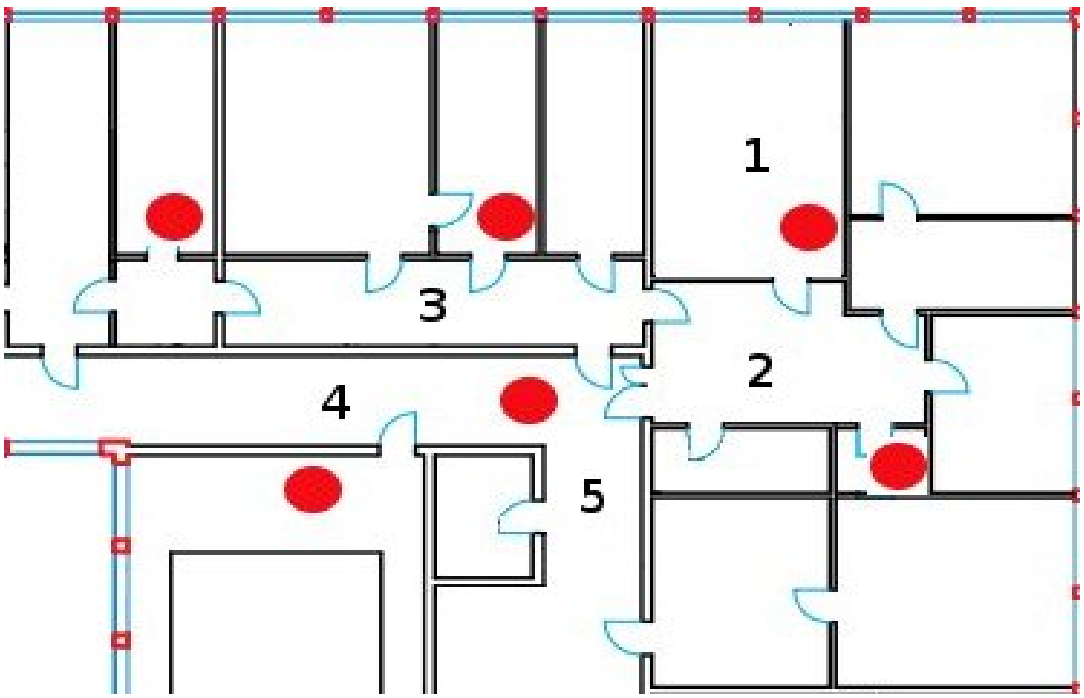
Agenda

1. Application Context
2. Method
   - Benchmark
   - Social Boosting
3. **Dataset**
4. Results
5. Summary

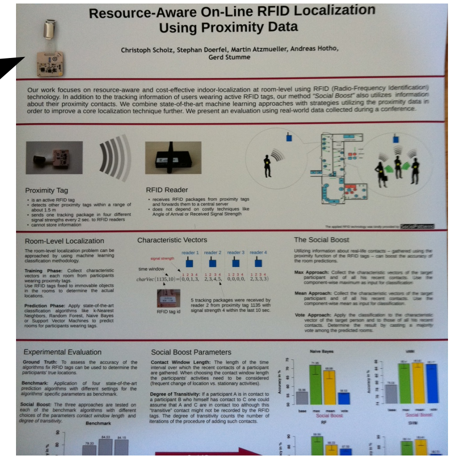


Data Collection

• real-time data collected during the two hour poster-session at the LWA 2010 in Kassel.
• we asked each participant to wear a proximity tag.
• 46 people took part in our experiment.
• we placed 6 RFID readers in the conference area.



Definition: Object Tag and User Tag



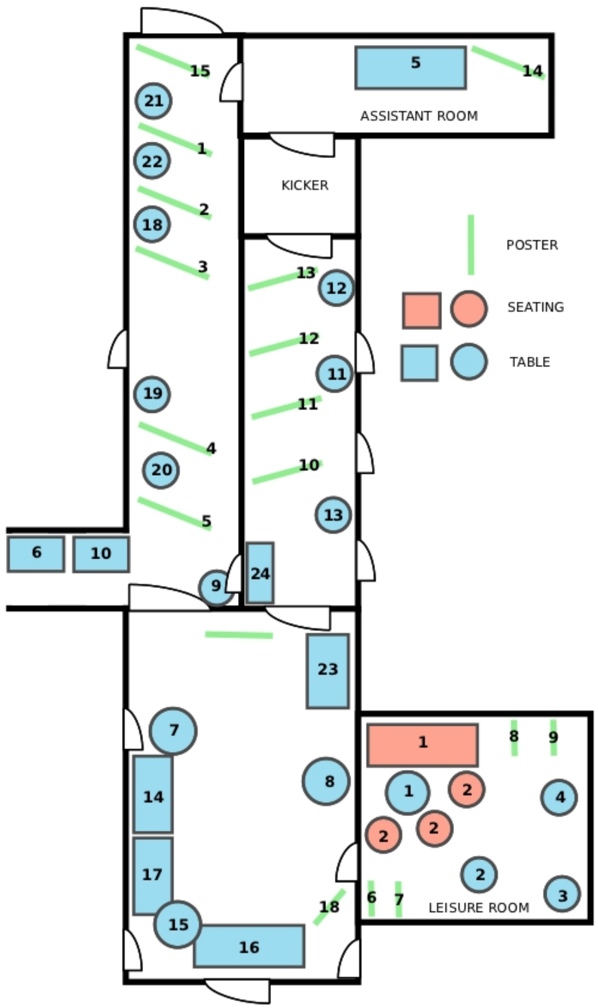
Object Tag

- is a proximity tag fixed to an unmovable object

User Tag

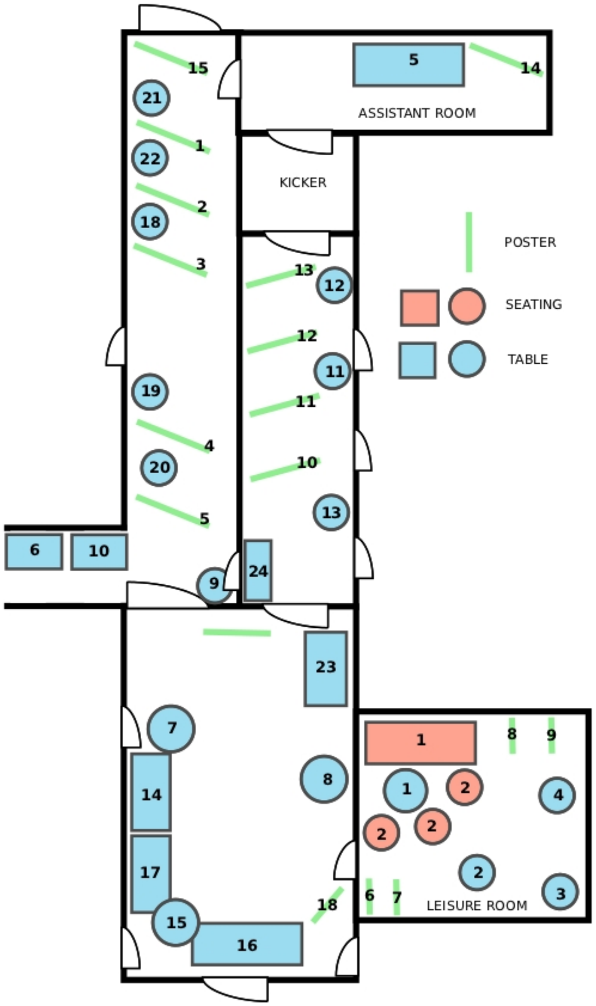
- is a proximity tag worn by a participant during the conference

Ground Truth



- we determined the position of the participants by using object tags
- whenever a participant's tag was recorded by an object tag, we know that the participant was in the same room as the object tag
- for the experiments we predicted the rooms for the vectors where the location could be verified.

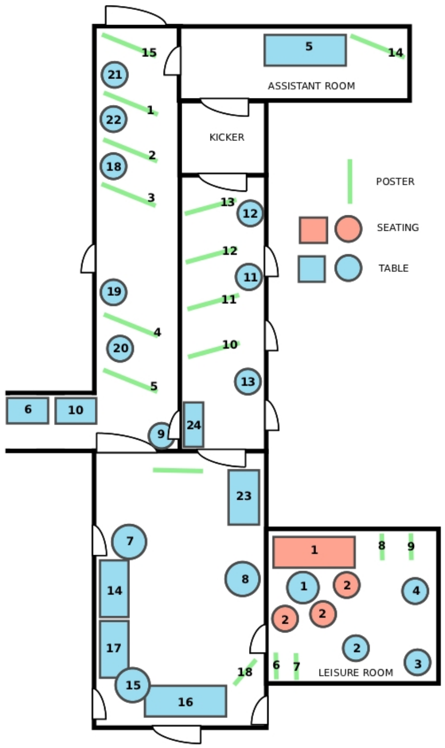
Training Data



- the first 1500 characteristic vectors collected through the object tags (ca. 20 min)

Agenda

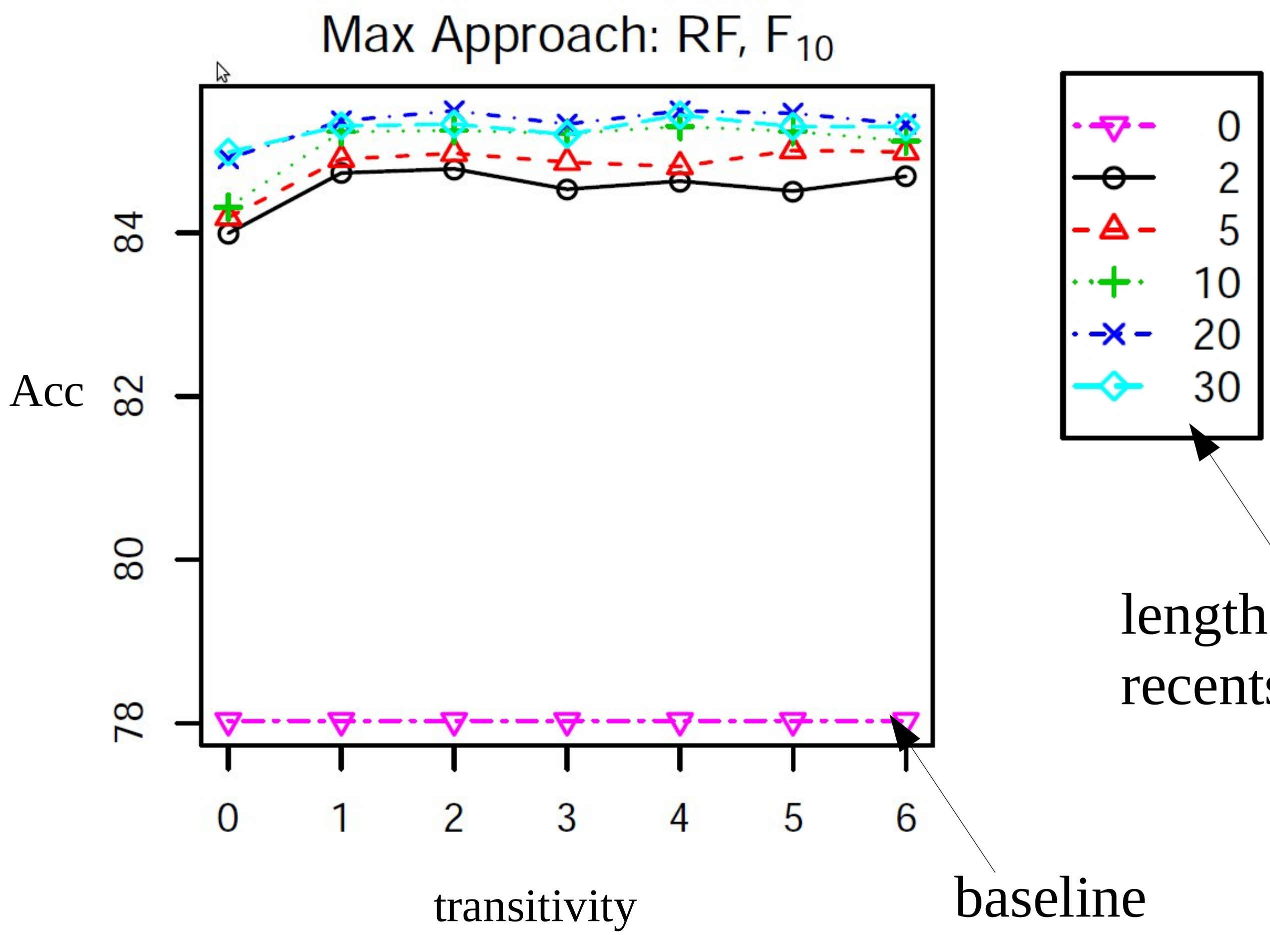
1. Application Context
2. Method
   - Benchmark
   - Social Boosting
3. Dataset
4. **Results**
5. Summary



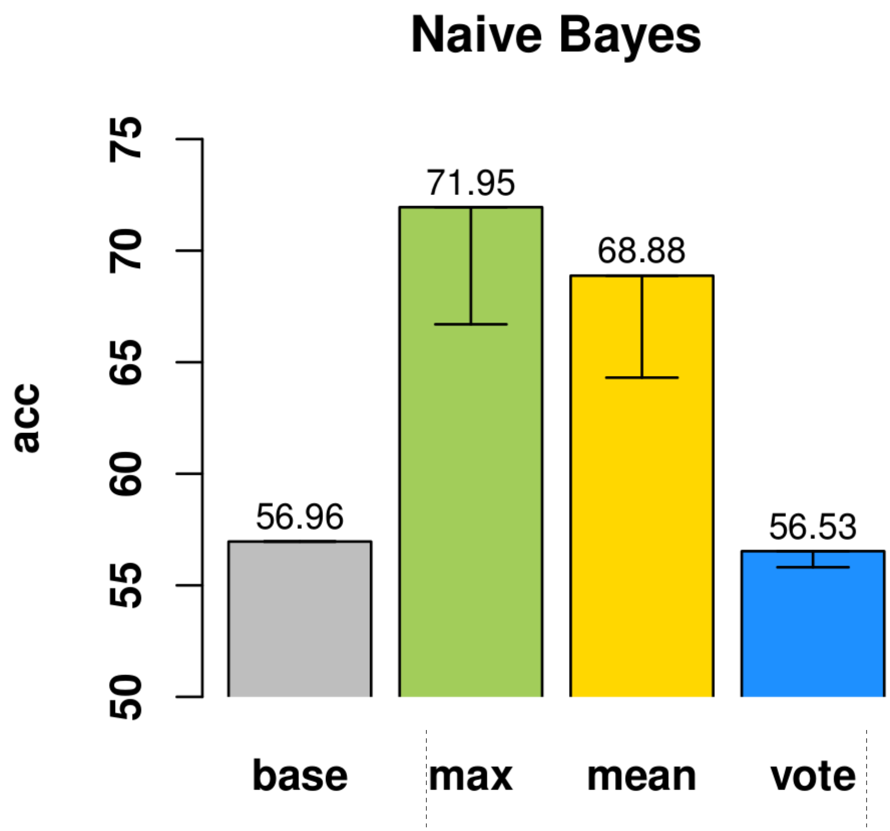
Parameters State of the Art Room Prediction

base		$l = 10$	$l = 30$	$l = 50$
kNN	k	165	185	200
	Accuracy	**73.58**	**79.33**	**79.80**
RF	$mtry$	1	2	4
	$ntree$	375	350	200
	Accuracy	**78.03**	**84.18**	**84.78**
SVM	j	1	1	1
	γ	-14	-18	-20
	Accuracy	**77.95**	**84.15**	**84.84**
nBay	**Accuracy**	**38.97**	**56.96**	**61.97**

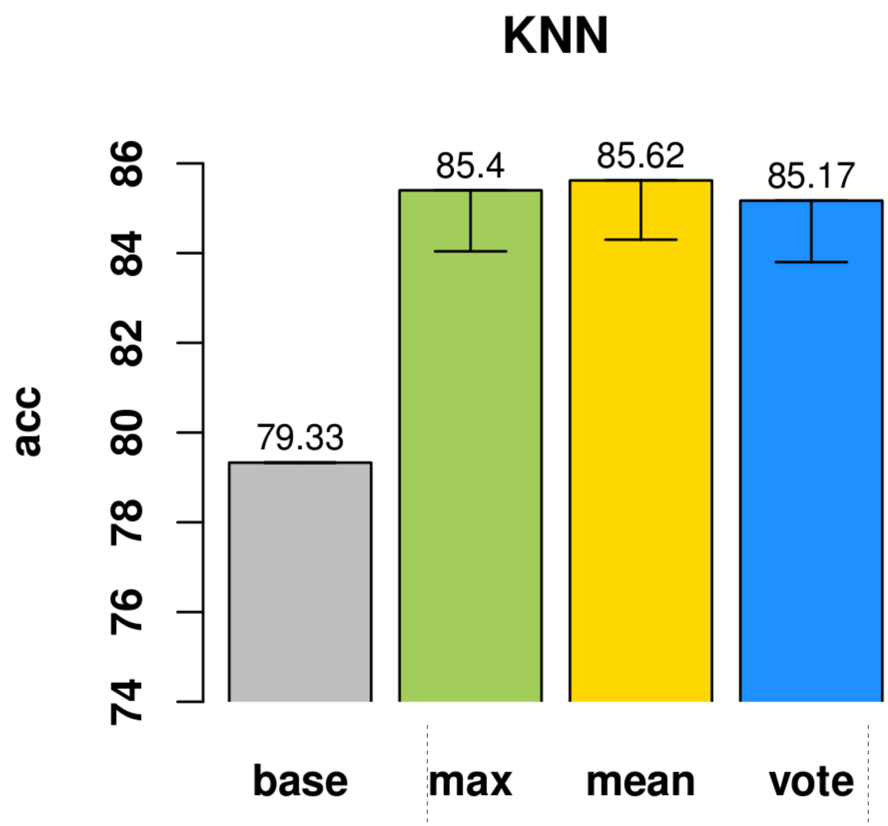
Parameters: Transitivity of Contacts



Social Boosting



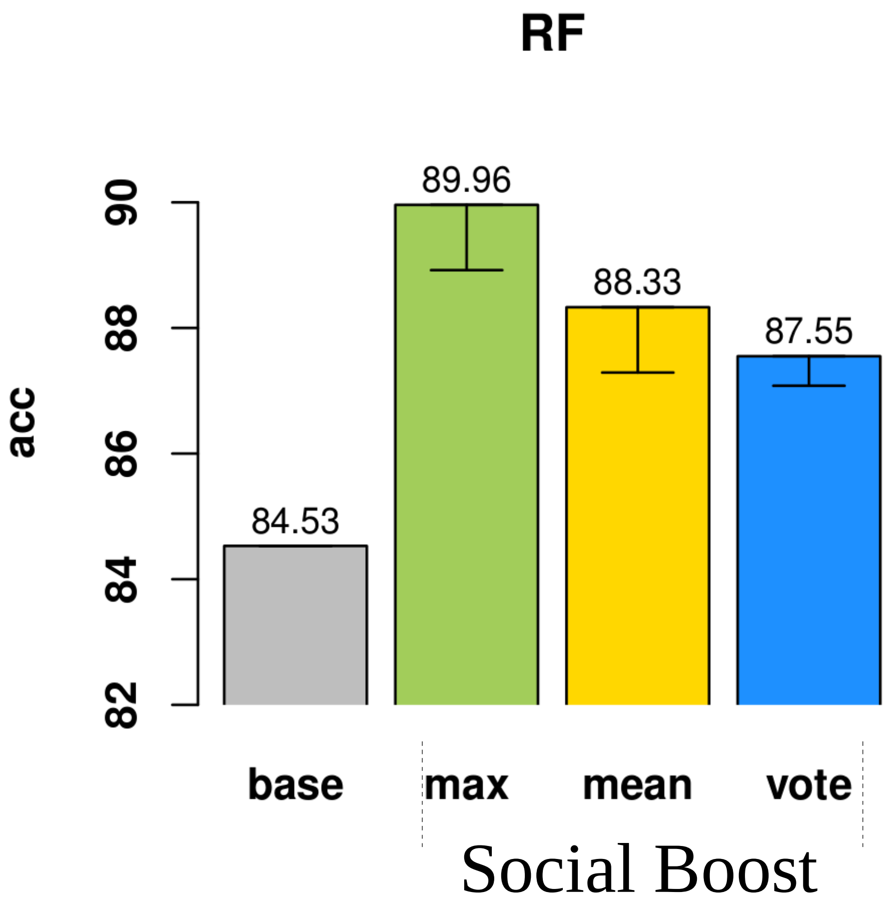
Social Boost

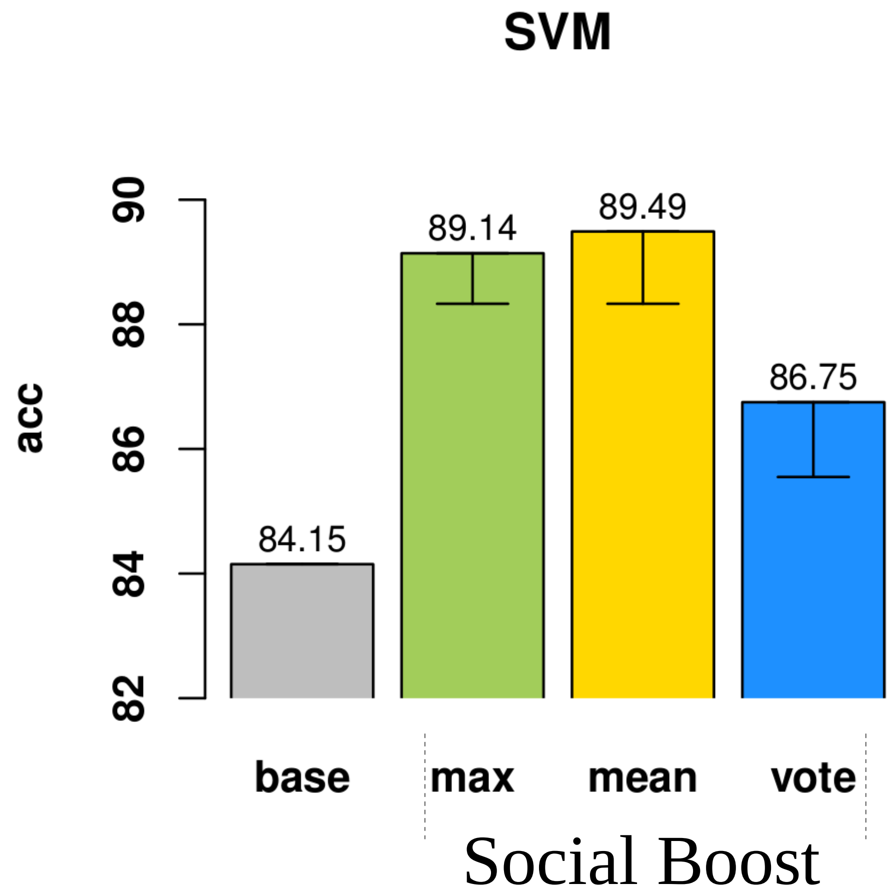


Social Boost

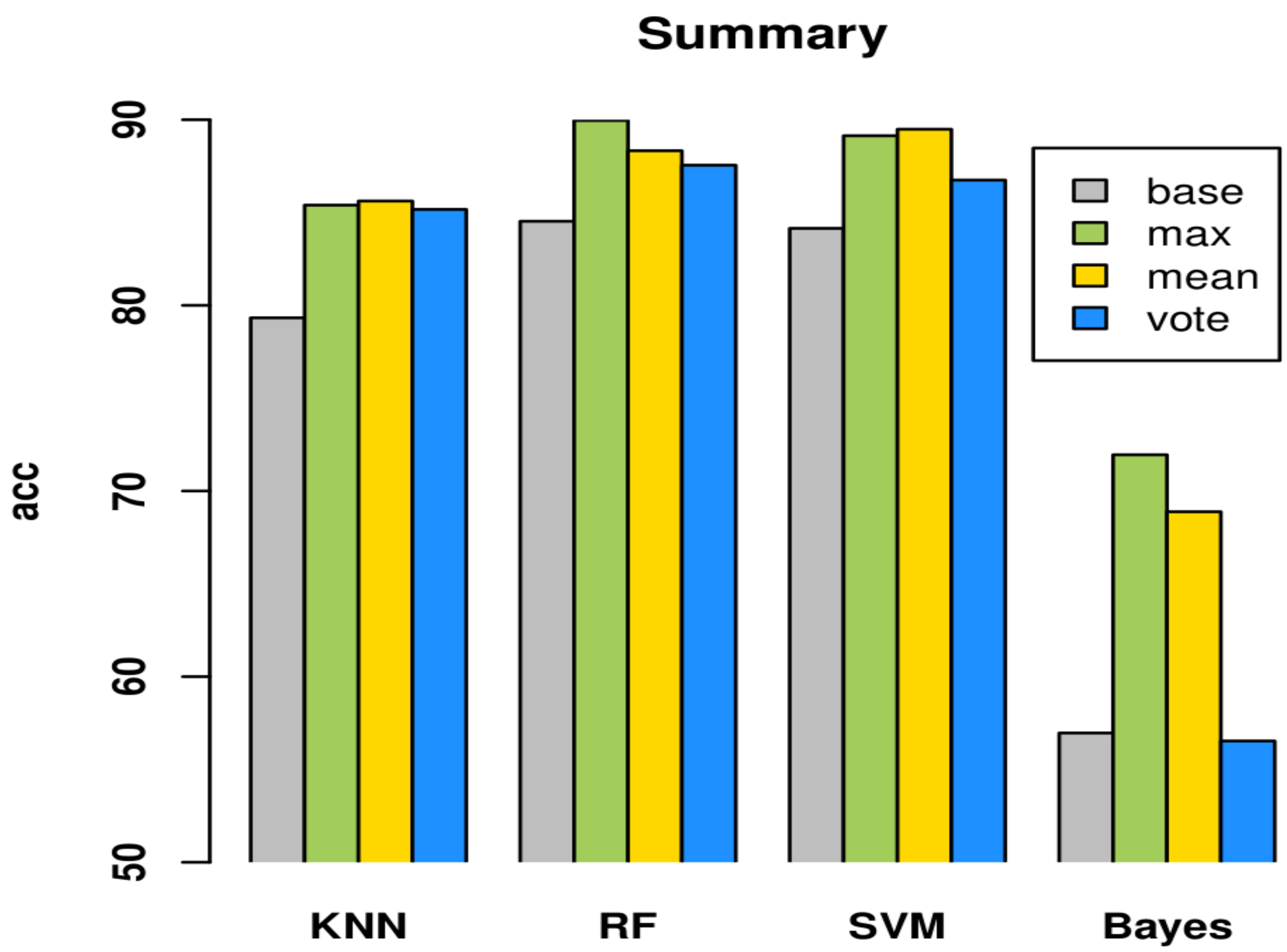
Scholz et al: Resource-Aware On-Line RFID Localization Using Proximity Tags, ECML 2011

Social Boosting





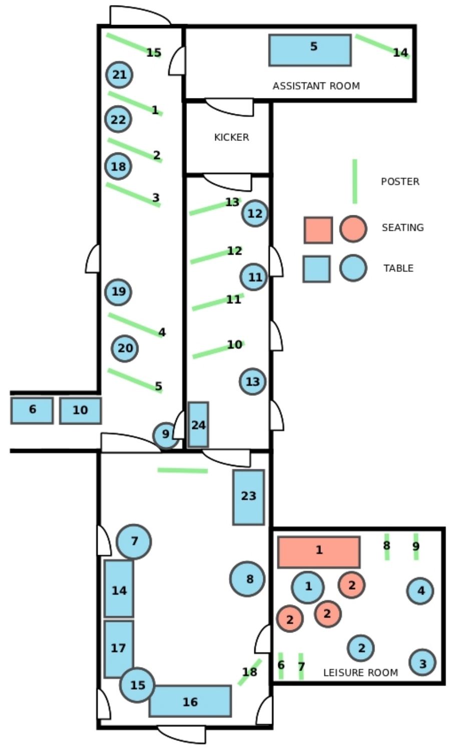
Social Boosting



Scholz et al: Resource-Aware On-Line RFID Localization Using Proximity Tags, ECML 2011

Agenda

1. Application Context
2. Method
   - Benchmark
   - Social Boosting
3. Dataset
4. Results
5. **Summary**



Summary and Future Work

- we presented methods that improve the localization accuracy significantly.
- RF and SVM perform very well, Naive Bayes not
- best result we got was 93 % with social boosting (only participants, who are in contact)

- learning of user data instead of object data
- using the localization history (HMM, CRFs,....)
- increasing the package frequency
- use additional datasets

Thanks for your attention.

- **BACKUP SLIDES**

Social Boosting, Boost

	kNN	RF	SVM	nBay
strategy	F_{50}	F_{10}	F_{10}	F_{10}
contact fraction	69.23	65.01	64.91	69.3
contact base ACC	74.33	76.83	77.78	33.05
contact best ACC	88.09	88.18	89.34	58.80
boost	13.76	11.34	11.57	25.74