清華大學
Tsinghua University

User Intention Inference

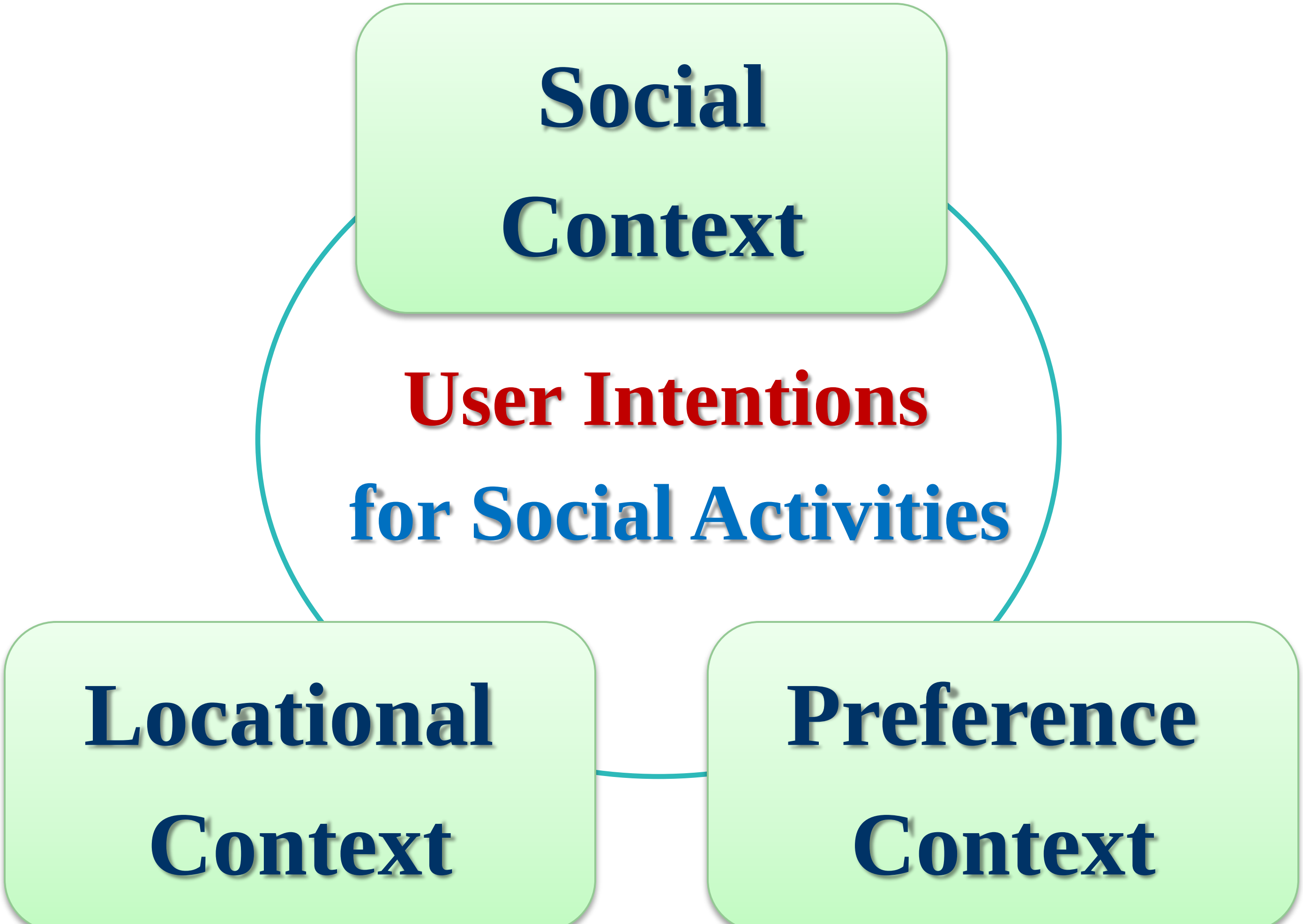
Tsinghua-BST Joint Project

PIs: Peng Cui, Yingjie Zhang, Shiqiang Yang

Computer Science & Technology Department

Tsinghua University

MOTIVATIONS



DELIVERABLES

Data

2443 activities

5643 comments

20+ activity categories

27 testees

Algorithm

Location-based

CF-based

Influence-based

Preference-based

Hybrid factor

System

Function completed

Robust and stable

Message encrypted

Data flow efficient

DATA

Data

2443 activities

5643 comments

20+ activity categories

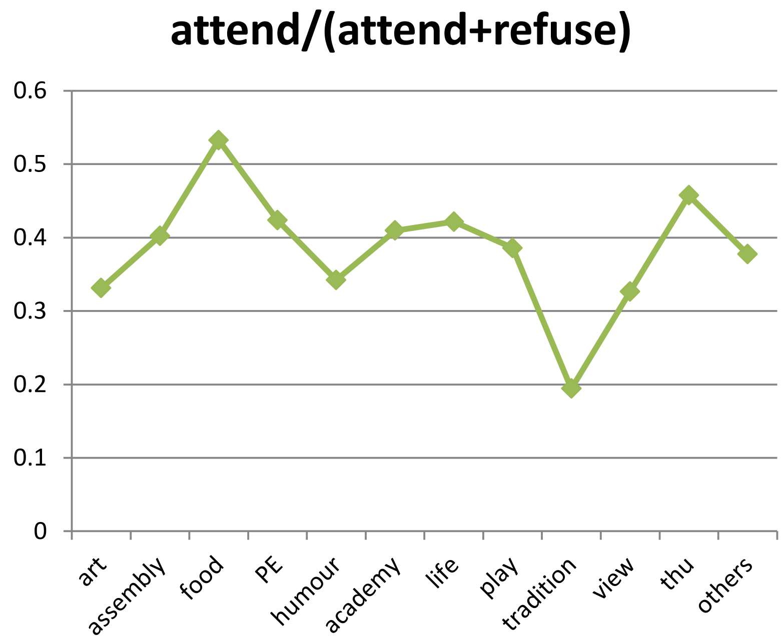
27 testees

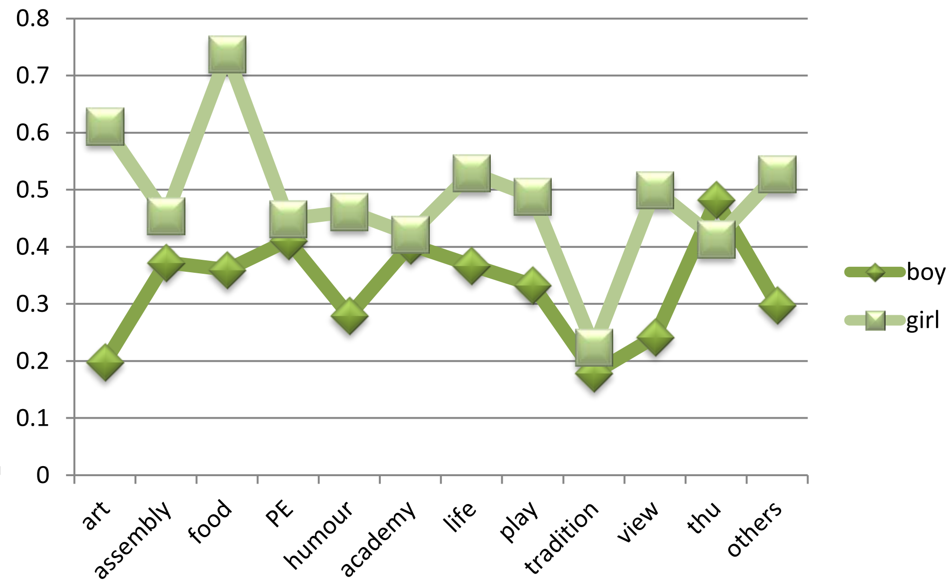
Content

Preference

Behavior

CONTENT





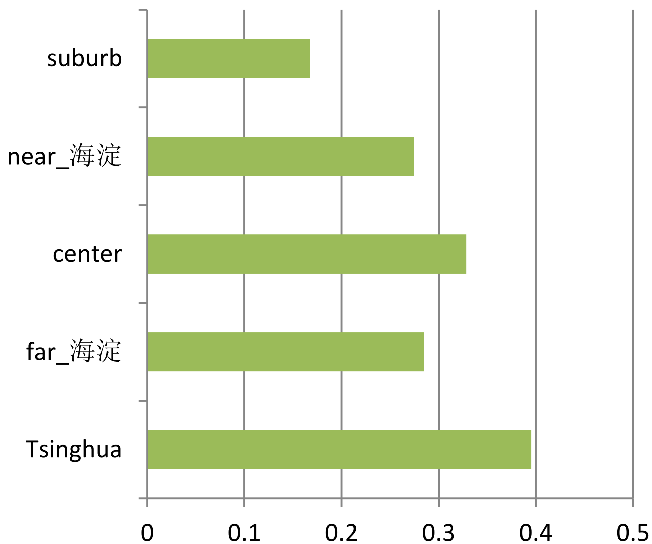
Conclusion

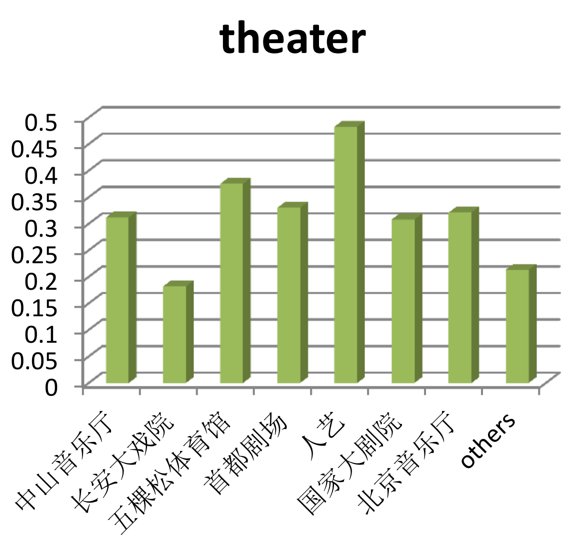
1, Students prefer activities in THU, which accounts for the common knowledge that people tend to choose familiar environment.

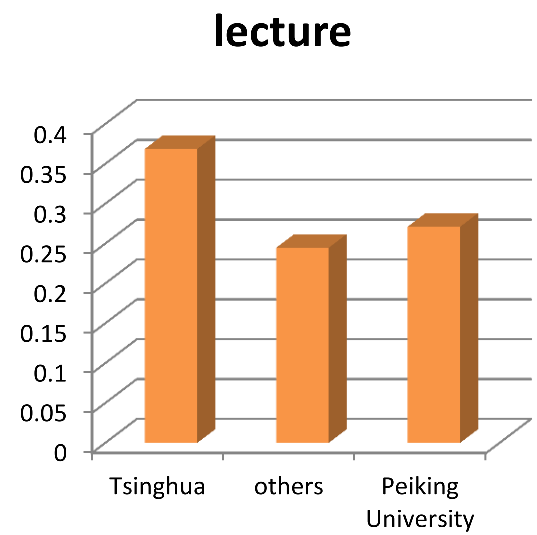
2, College students are more active in outside activities which is reflected in food and play peak.

3, Girls are more active in attending different kinds of activities.

CONTENT (cont'd)







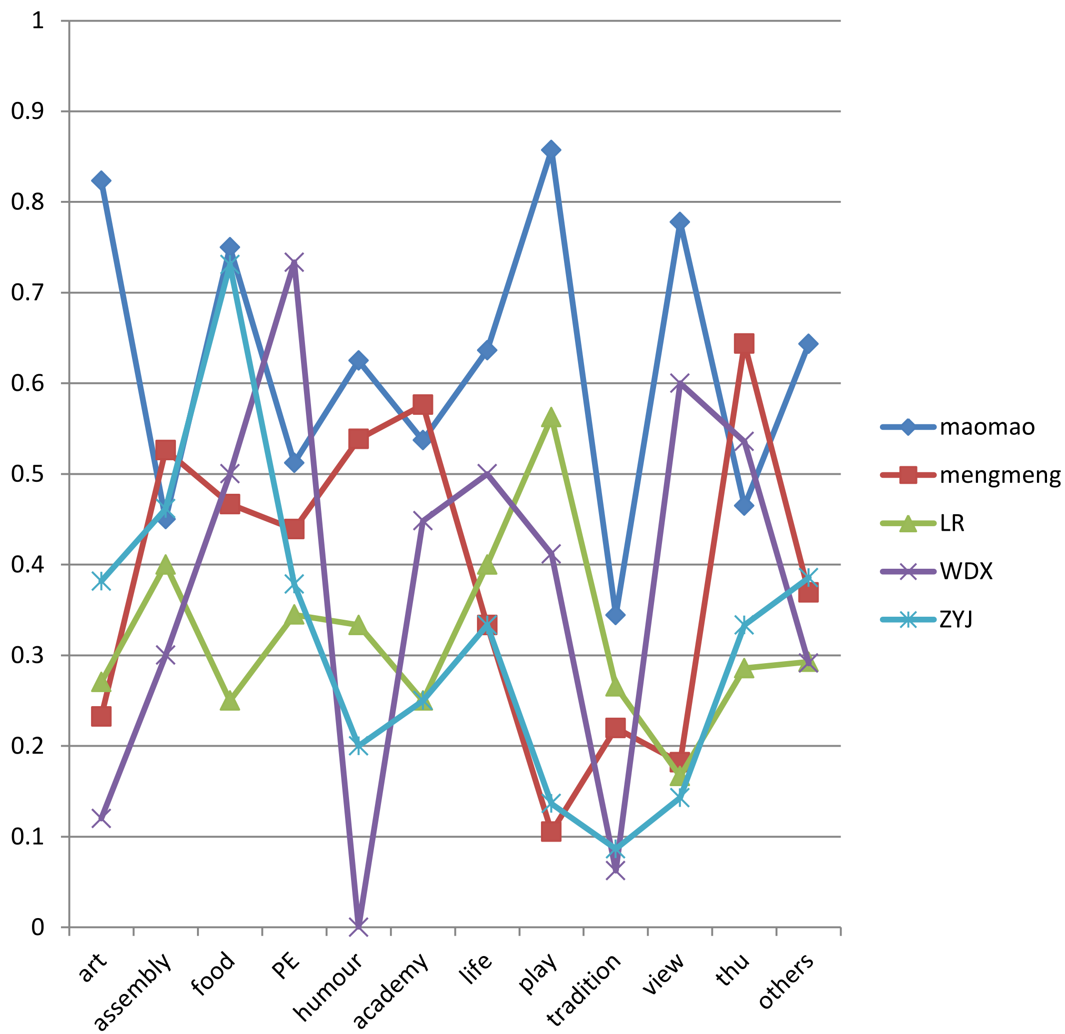
Conclusion

1, people sometimes prefer activities near their living or working places because of the less cost in distance.

2, but this is not always the case, not even a salient measure.

3, we can't use location as a simple rule for heuristic usage.

PREFERENCE

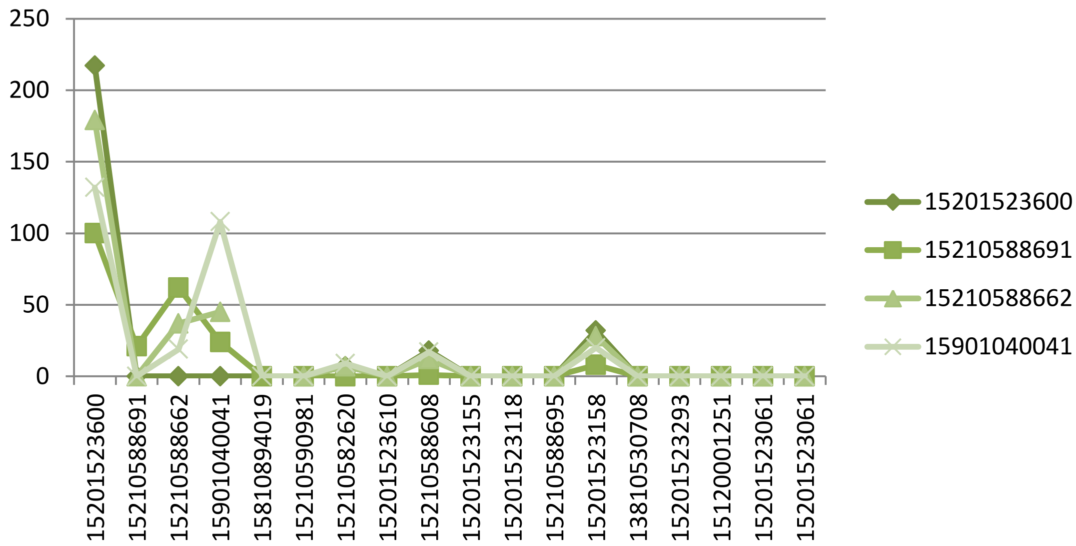


Conclusion

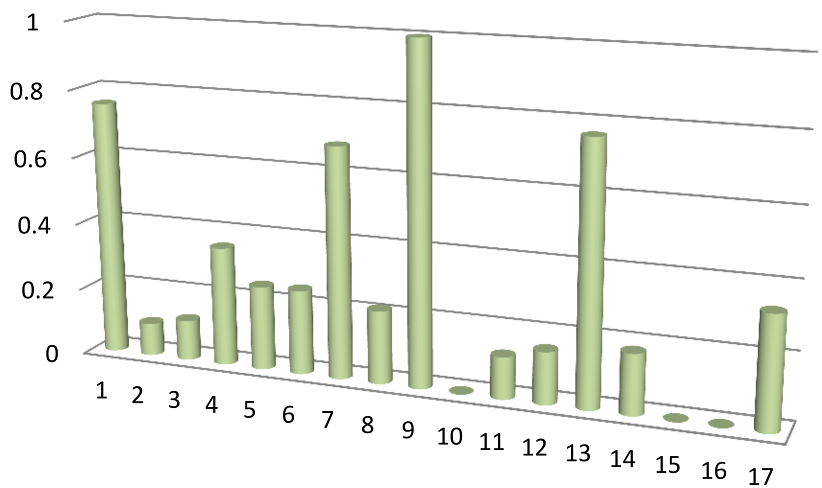
1, Different people has distinct predilection on activity topic choice.

2,People's comments, attend percent or refuse percent has a connection with their personal characteristics.

BEHAVIOR



ratio between preference and influence



Conclusion

1, The more activities a person takes part in, the more influence he/she has on others.

2, Different people are widely divergent in facing social influence.

3, In most cases, outgoing people would like to choose according to their personal preference.

ALGORITHM

Algorithm

Location-based

CF-based

Influence-based

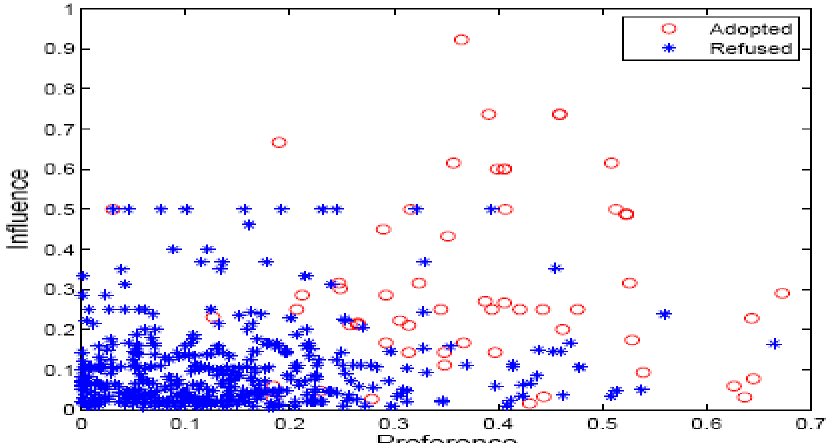
Preference-based

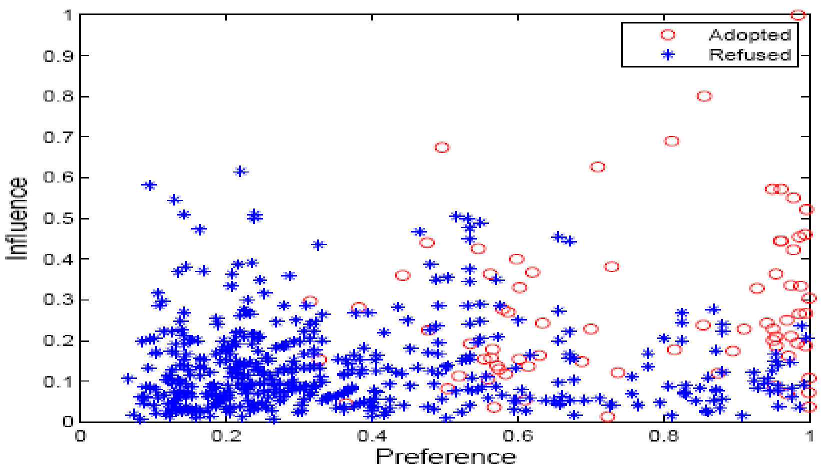
Hybrid factor

Preliminary | Algorithm | Evaluation

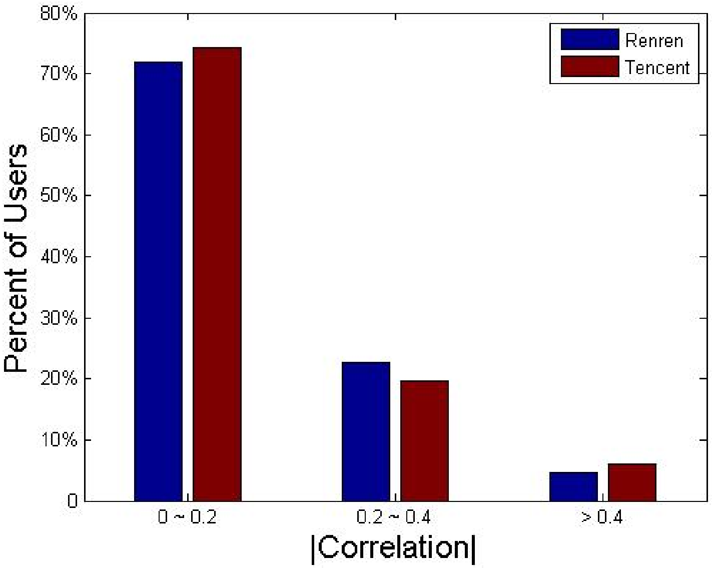
PRELIMINARY

Accepted > Refused on Preference and Influence



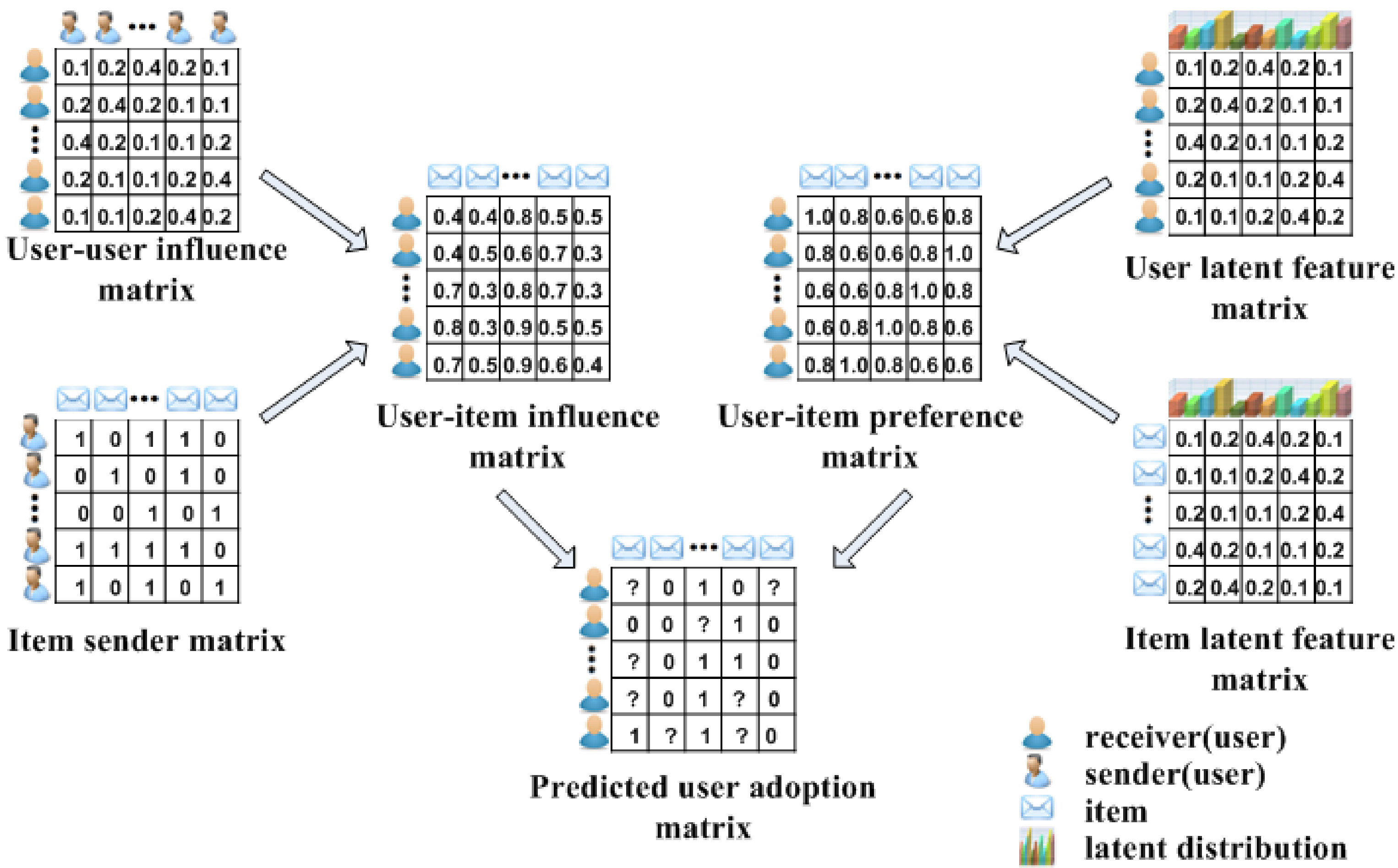


No Correlation between Preference and Influence



PRELIMINARY

Framework



ALGORITHM

Social Contextual Recommendation

$$P(\mathbf{R}|\mathbf{S},\mathbf{U},\mathbf{V},\sigma_R^2)=\prod_{i=1}^{M}\prod_{j=1}^{N}\mathcal{N}(\mathbf{R}_{ij}|\mathbf{S}_i\mathbf{G}_j^\top\odot\mathbf{U}_i^\top\mathbf{V}_j,\sigma_R^2)$$

$$\mathcal{J}=||\mathbf{R}-\mathbf{S}\mathbf{G}^\top\odot\mathbf{U}^\top\mathbf{V}||_F+\alpha||\mathbf{W}-\mathbf{U}^\top\mathbf{U}||_F$$
$$+\beta||\mathbf{C}-\mathbf{V}^\top\mathbf{V}||_F+\gamma||\mathbf{S}-\mathbf{F}||_F$$
$$+\delta||\mathbf{S}||_F+\eta||\mathbf{U}||_F+\lambda||\mathbf{V}||_F$$

$$\frac{\partial\mathcal{J}}{\partial\mathbf{S}}=2\Big(-\mathbf{R}(\mathbf{G}\odot\mathbf{V}^\top\mathbf{U})+(\mathbf{S}\mathbf{G}^\top\odot\mathbf{U}^\top\mathbf{V})\mathbf{G}+\gamma(\mathbf{S}-\mathbf{F})+\delta\mathbf{S}\Big)$$

$$\frac{\partial\mathcal{J}}{\partial\mathbf{U}}=2\Big(-\mathbf{V}\mathbf{R}^\top+\mathbf{V}(\mathbf{G}\mathbf{S}^\top\odot\mathbf{V}^\top\mathbf{U})-2\alpha\mathbf{U}\mathbf{W}+2\alpha\mathbf{U}\mathbf{U}^\top\mathbf{U}+\eta\mathbf{U}\Big)$$

$$\frac{\partial\mathcal{J}}{\partial\mathbf{V}}=2\Big(-\mathbf{U}\mathbf{R}+\mathbf{U}(\mathbf{S}\mathbf{G}^\top\odot\mathbf{U}^\top\mathbf{V})-2\beta\mathbf{V}\mathbf{C}+2\beta\mathbf{V}\mathbf{V}^\top\mathbf{V}+\lambda\mathbf{V}\Big)$$

Algorithm 1 Social Contextual Model Gradient Algorithm

Require: $0<\alpha_S^{(k)},\alpha_U^{(k)},\alpha_V^{(k)}<1$, $k=0$. Initialization $\mathcal{J}^{(0)}=\mathcal{J}(\mathbf{S}^{(0)},\mathbf{U}^{(0)},\mathbf{V}^{(0)})$.

Ensure: $\mathcal{J}^{(0)}\geq 0$, $\mathcal{J}^{(k+1)}<\mathcal{J}^{(k)}$

for $k=1,2,\cdots$ **do**

Calculate $\frac{\partial\mathcal{J}}{\partial\mathbf{S}}^{(k-1)}$, $\frac{\partial\mathcal{J}}{\partial\mathbf{U}}^{(k-1)}$, $\frac{\partial\mathcal{J}}{\partial\mathbf{V}}^{(k-1)}$

$\mathbf{S}^{(k)}=\mathbf{S}^{(k-1)}-\alpha_S^{(k-1)}\cdot\frac{\partial\mathcal{J}}{\partial\mathbf{S}}^{(k-1)}$

$$\mathcal{J}^{(k)}\leftarrow\mathcal{J}(\mathbf{S}^{(k)},\mathbf{U}^{(k-1)},\mathbf{V}^{(k-1)})$$

$\mathbf{U}^{(k)}=\mathbf{U}^{(k-1)}-\alpha_U^{(k-1)}\cdot\frac{\partial\mathcal{J}}{\partial\mathbf{U}}^{(k-1)}$

$$\mathcal{J}^{(k)}\leftarrow\mathcal{J}(\mathbf{S}^{(k)},\mathbf{U}^{(k)},\mathbf{V}^{(k-1)})$$

$\mathbf{V}^{(k)}=\mathbf{V}^{(k-1)}-\alpha_V^{(k-1)}\cdot\frac{\partial\mathcal{J}}{\partial\mathbf{V}}^{(k-1)}$

$$\mathcal{J}^{(k)}\leftarrow\mathcal{J}(\mathbf{S}^{(k)},\mathbf{U}^{(k)},\mathbf{V}^{(k)})$$

end for

ALGORITHM FOR APPLICATION

Complexity

Space $O(M(M+N))$

Time $O(N^3T)$

Practical

Linear Model

Space $O(MN)$

Time $O(MN)$

$val(u,a) = \alpha \cdot val_{itemCF}(u,a) + \beta \cdot val_{preference}(u,a) + \gamma \cdot val_{influence}(u,a) + \delta \cdot val_{location}(u,a), \alpha, \beta, \gamma, \delta \epsilon (0,1)$

Hybrid： Item-based CF + Preference-based + Influence-based + Location-based Recommendation

Matrix Factorization

$$\mathbf{R} \leftarrow \mathbf{S}\,\mathbf{G}^{\mathrm{T}} \odot \mathbf{U}\,\mathbf{V}^{\mathrm{T}}$$

EVALUATION

MAE (Mean Absolute Error), RMSE (Root Mean Square Error)

The lower, the better. [0,1]

Kendall, Spearman (Ranking-based)

The higher, the better. [-1, 1]

Prec@k (Precision at top-k)

The higher, the better. [0, 100%]

	MAE	RMSE	Kendall	Spearman	Prec @5	Prec @10	Prec @15	Prec @20
Latest	0.995	1.148	0.006	0.008	40.4%	40.4%	40.7%	40.7%
Location	0.695	0.987	0.079	0.117	52.1%	50.3%	49.2%	48.7%
Hybrid	0.664	0.805	0.245	0.361	**83.3%**	**80.5%**	**77.2%**	**74.5%**

ALGORITHMIC TOOLSET

Location-based

User likes activities that are around him.

Mobile services. Restaurant, Café.

For user u and activity a,

$$val(u,a) = val\big(u, place(a)\big) = \sum_{a' \in Activities(u)} val(u, place(a'))$$

CF-based

User likes activities that his friends like.

E-commerce. Movie, music, phone, restaurant.

For user u and activity a,

$$val(u,a) = \sum_{v \in Friend(u)} val(v,a)$$

ALGORITHMIC TOOLSET

Influence-based

User likes activities from close relationships.

Social networks. Microblogging.

For user u and activity a,

$$val(u,a) = val\big(u, sender(a)\big) = \sum_{a\prime \in Activities(u)} val(u, sender(a'))$$

Preference-based

User likes activities that he always likes.

Web portal. News.

For user u and activity a,

$$val(u,a) = \sum_{a\prime \in Activities(u)} val(u, a')$$

DELIVERABLES

System

Function completed

Robust and stable

Message encrypted

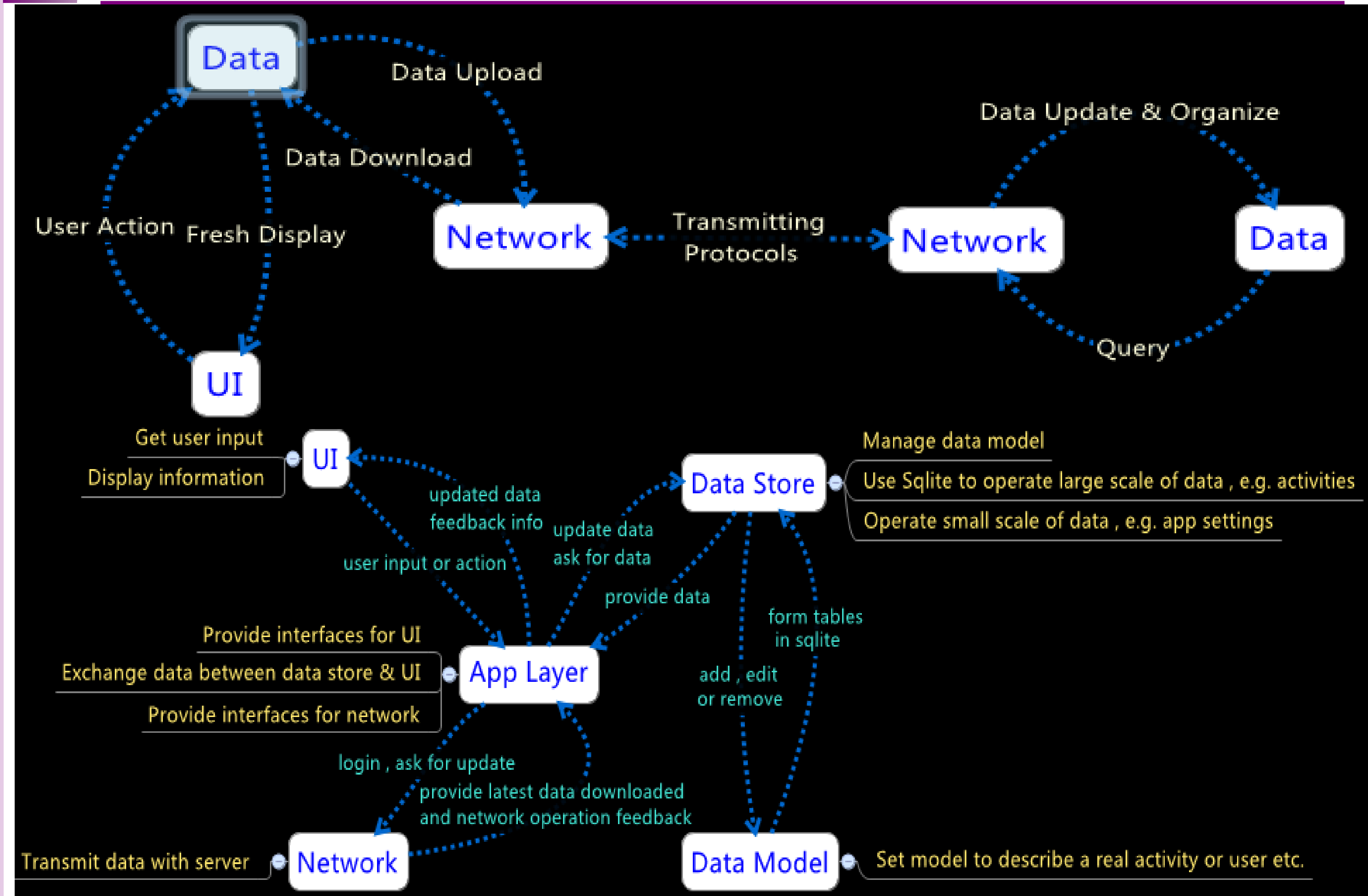
Data flow efficient

Framework

Stability

Flow Efficiency

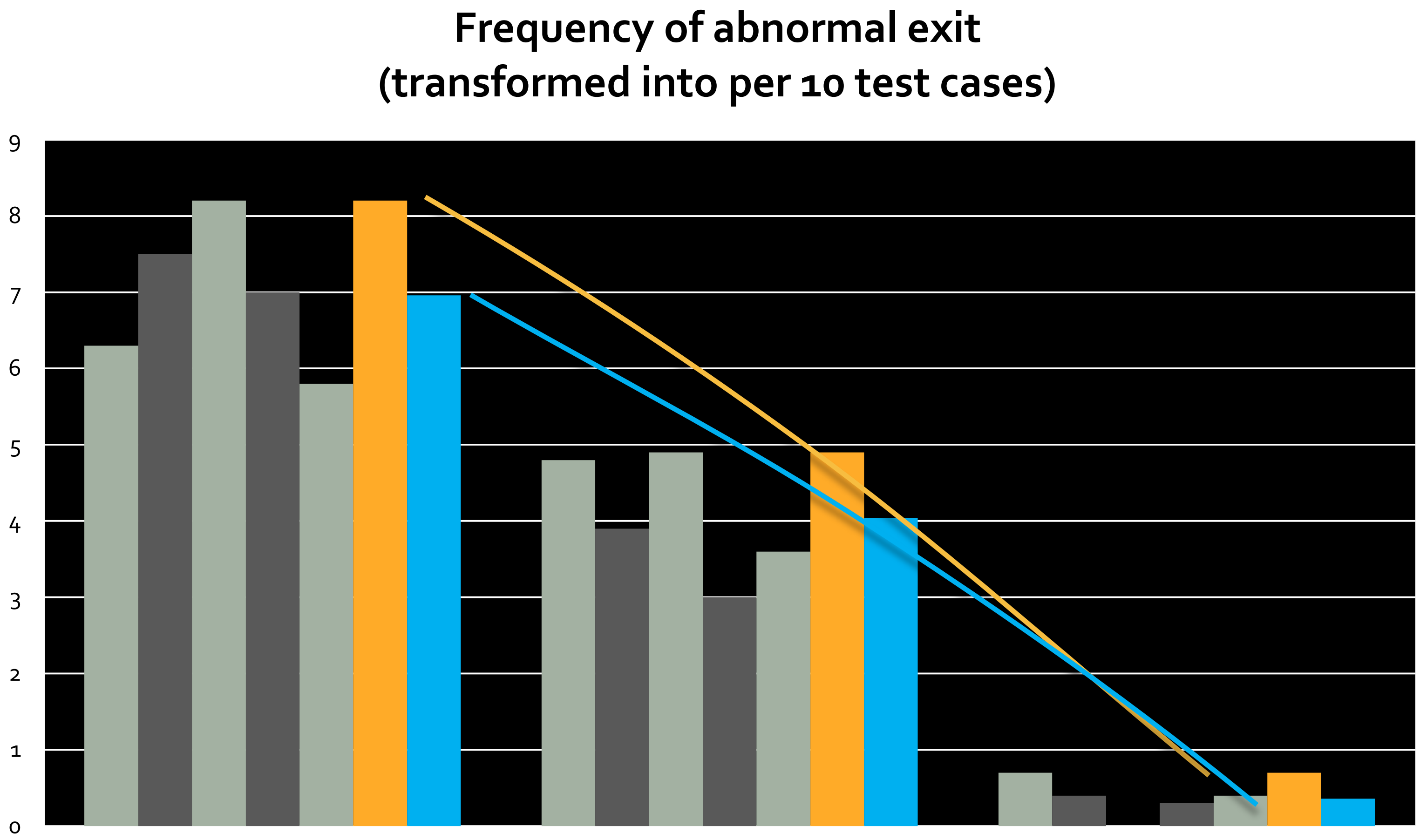
FRAMEWORK



Data
Data Upload
Data Download
User Action
Fresh Display
Network
Transmitting
Protocols
Network
Data Update & Organize
Data
Query
UI
Get user input
Display information
UI
updated data
feedback info
user input or action
update data
ask for data
Data Store
Manage data model
Use Sqlite to operate large scale of data , e.g. activities
Operate small scale of data , e.g. app settings
provide data
form tables
in sqlite
Provide interfaces for UI
Exchange data between data store & UI
Provide interfaces for network
App Layer
add , edit
or remove
login , ask for update
provide latest data downloaded
and network operation feedback
Transmit data with server
Network
Data Model
Set model to describe a real activity or user etc.

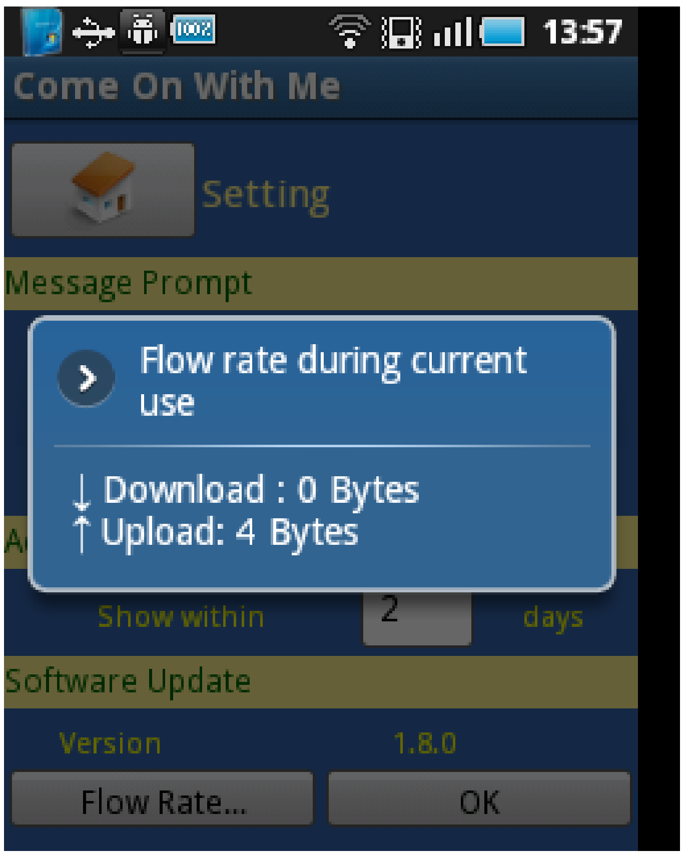
STABILITY

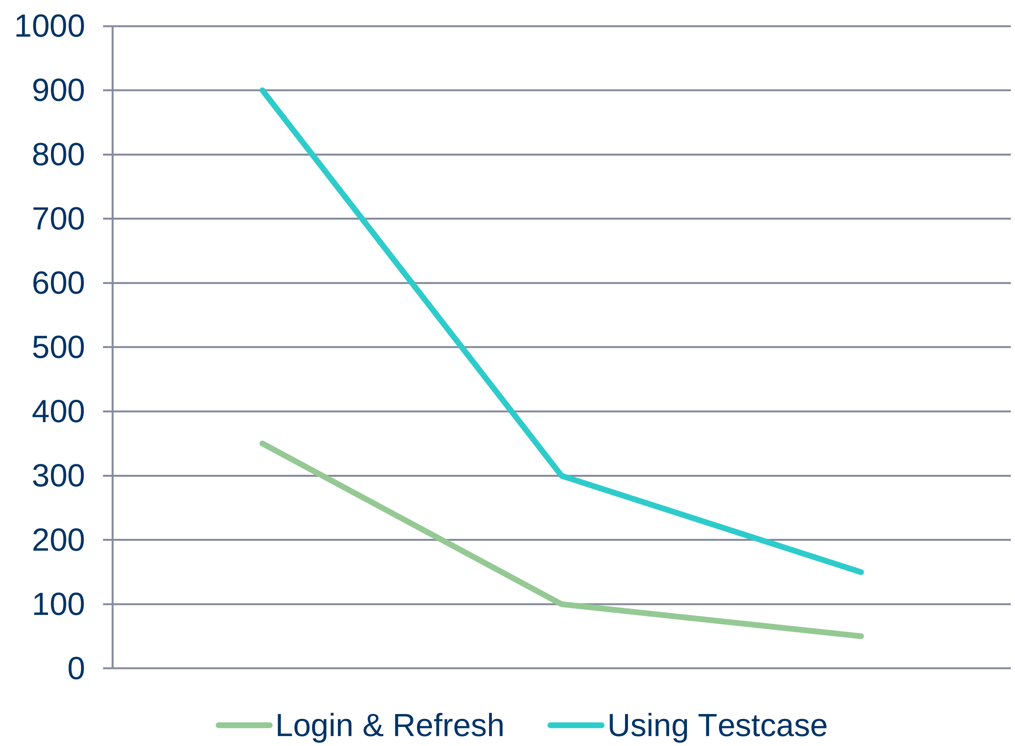
- The complex communication between server and client, and dynamic and intensive user behaviors make the system unstable.
- Now the system satisfy the requirement of practical use.



FLOW EFFICIENCY

- The complex communication between server and client, and dynamic and intensive user behaviors make the system unstable.
- Now the system satisfy the requirement of practical use.





CONCLUSIONS

- The collected dataset is adequate to support following research:
  - User Modeling
  - Social Ties Inference
  - Intelligent Recommendation
- The algorithmic toolset includes:
  - Location-based Recommendation
  - Preference-based Recommendation
  - Collaborative Filtering
  - Influence-based Recommendation
  - Hybrid Factor Recommendation
- The system is adequate for further data collection and inner testing.

THANKS

Merry Christmas

and

Happy New Year!