Leveraging Distributed Human Computation and Consensus Partition for Entity Coreference

Saisai Gong, Wei Hu, Yuzhong Qu
Websoft Research Group, Nanjing University, China

Contents

- Introduction
- Overview of approach
- Consensus partition
- Ensemble learning
- Evaluation
- Conclusion

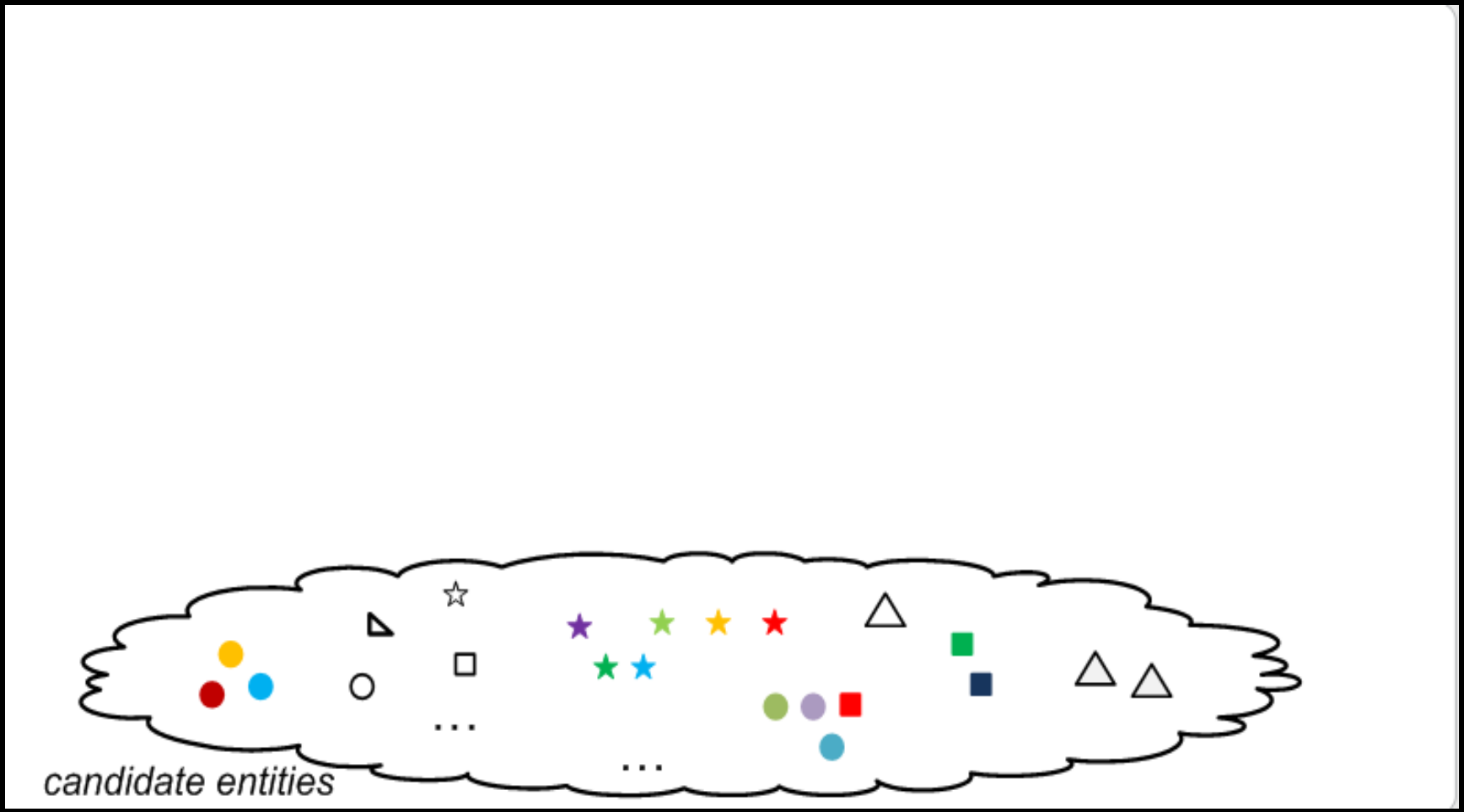
Introduction

- Human knowledge is valuable for entity coreference
- Crowdsourcing and other systems used for acquiring human contribution to coreference
- Quality of user-judged results is a problem
  - Mistakes and outliers
  - Omissions
  - …

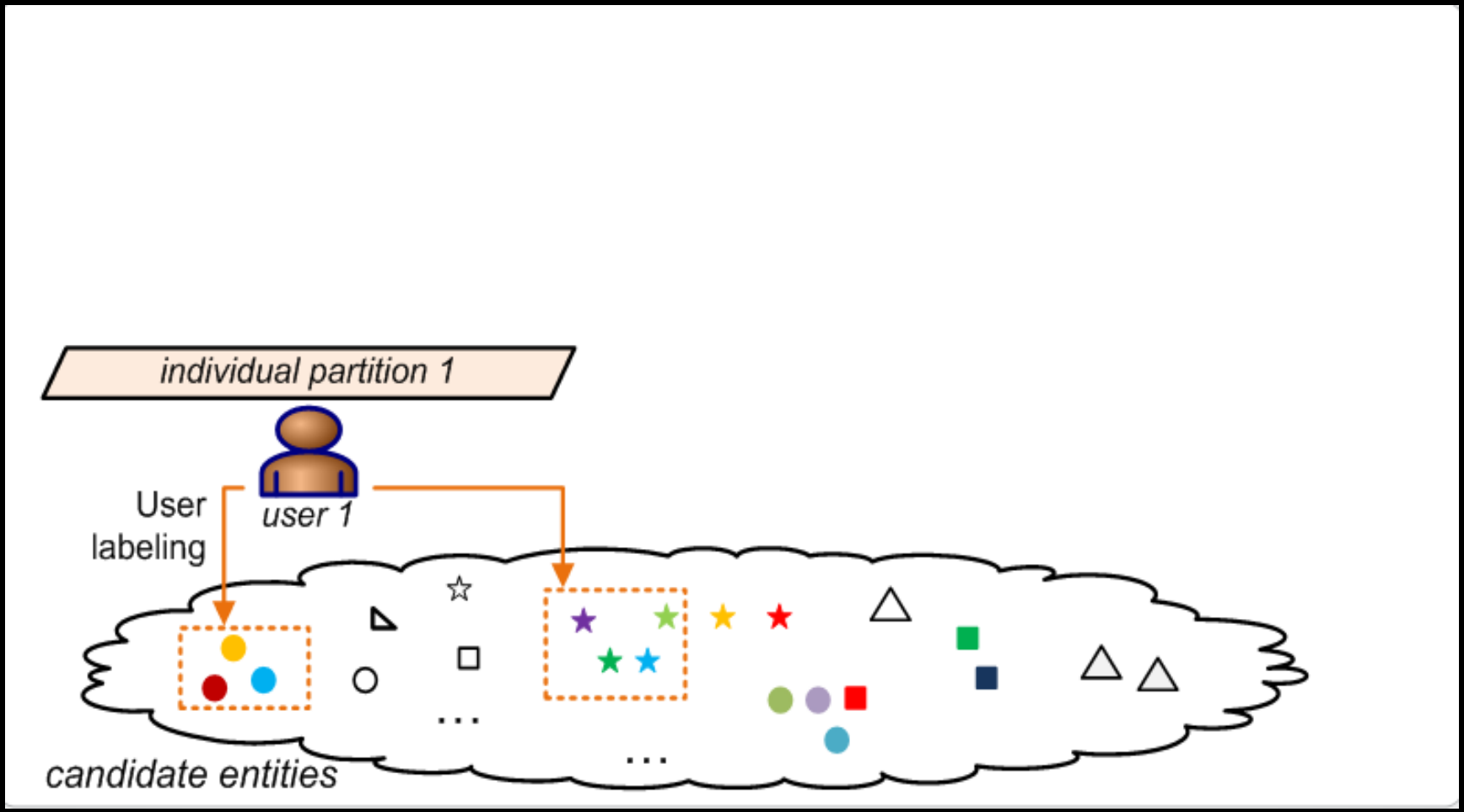
Introduction

- In this paper, we propose coCoref which
  - Acquire human contribution from web browsing activities
  - Improve quality by aggregating individual results using consensus partition
    - Consensus partition: minimize disagreements among partitions to obtain more robust and comprehensive results
  - Alleviate user involvement by automatically identifying coreferent entities not judged

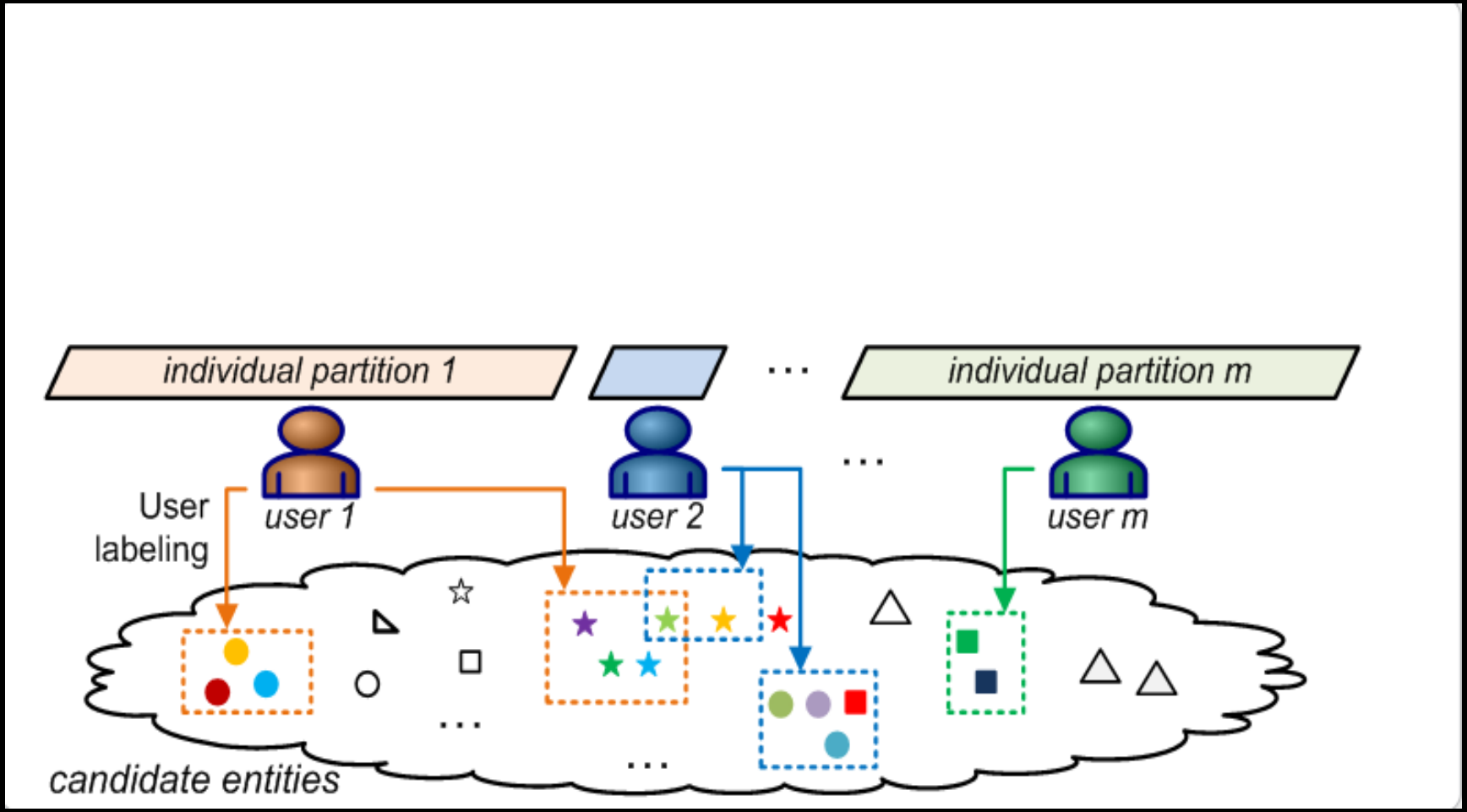
Overview of approach



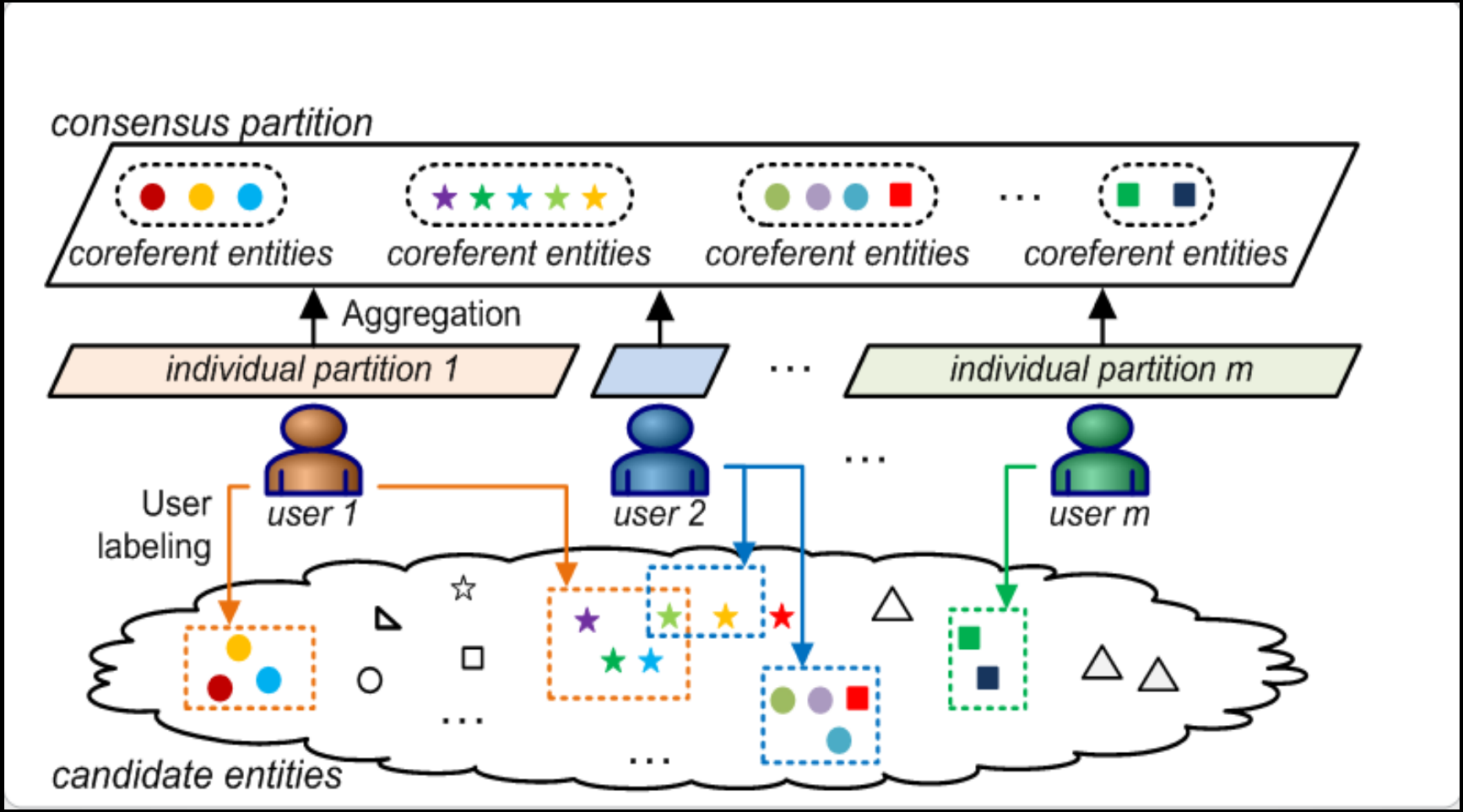
Overview of approach



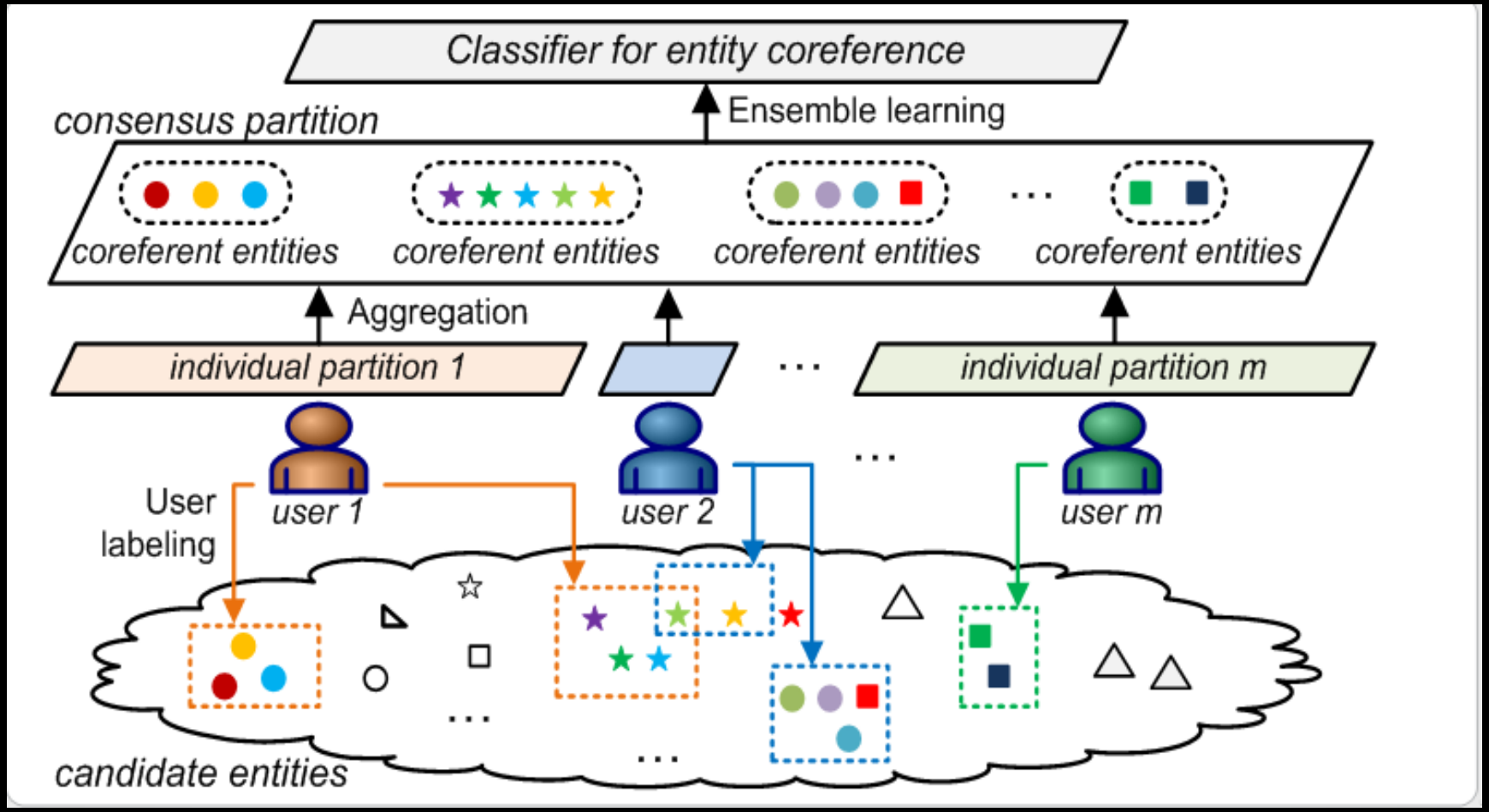
Overview of approach



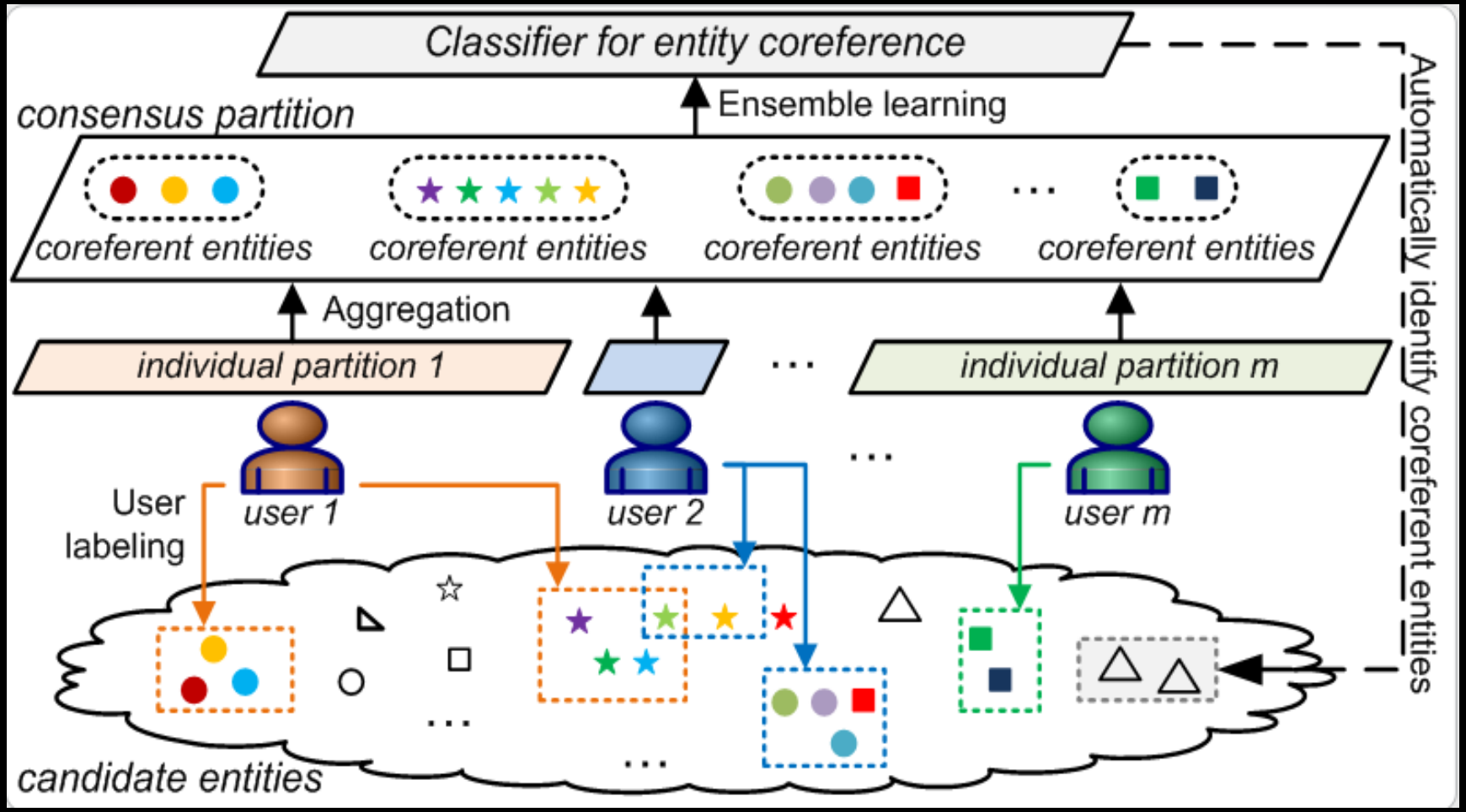
Overview of approach



Overview of approach



Overview of approach



Example


NewYorkCity, NY, TheBigApple
Alice
Manhattan, NewYorkCounty
NewYorkCity, TheBigApple, NewYorkCounty
Tom
NY, Manhattan
NewYorkCity, TheBigApple
Mike
NY, Manhattan

Example

Consensus partition

NewYorkCounty

NewYorkCity, TheBigApple

NY, Manhattan

Consensus partition

- Compute a consensus partition that minimize the number of disagreements with input partitions
  - Measure disagreement between partitions using symmetric difference distance

$$Dist_{\mathrm{T}}(\tau) = \sum_{\pi_j \in \mathrm{T}} Dist(\tau, \pi_j)$$
$$= \sum_{\pi_j \in \mathrm{T}} \sum_{v<w} (\delta_\tau(e_v, e_w)\psi_{\pi_j}(e_v, e_w) + (1 - \delta_\tau(e_v, e_w))\delta_{\pi_j}(e_v, e_w)), \quad (1)$$

- NP-complete

Consensus partition

- Approximation Algorithm

Algorithm 1: CC-Pivot [1]

Input: entity set $\mathbf{X}$, individual partition set T
Output: consensus partition τ on $\mathbf{X}$

Choose a pivot entity $e_v \in \mathbf{X}$ uniformly at random;
Let $C \leftarrow \{e_v\}$, $\mathbf{X}' \leftarrow \emptyset$;
foreach $e_w \in \mathbf{X}, w \neq v$ **do**
    if $\phi'_{vw} > 0$ **then**
        $C \leftarrow C \cup \{e_w\}$;
    else
        $\mathbf{X}' \leftarrow \mathbf{X}' \cup \{e_w\}$;
return $\tau \leftarrow \{C\} \cup \text{CC-Pivot}(\mathbf{X}', \mathrm{T})$;

3-approximation algorithm
time complexity: $\mathcal{O}(|\tau| \cdot |\mathbf{X}| \cdot |\mathrm{T}|)$

Ensemble learning

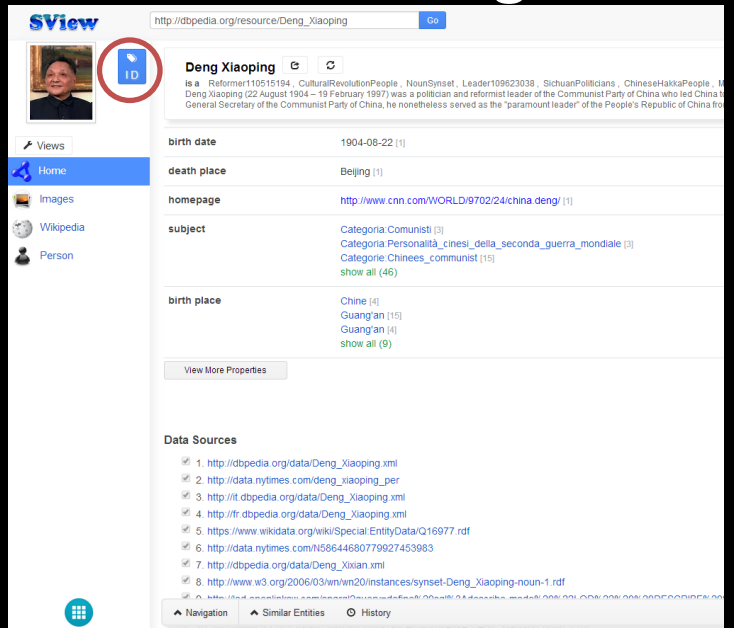
- Given training examples, learn a classifier to identify whether two entities are coreferent
- Training examples from consensus partition
  - Positive: entity pairs in the same equivalence class
  - Negative: entity pairs across different equivalence classes
- Feature: property-value similarity
- All property pairs for feature vector

Ensemble learning

- Choose suitable learning algorithms when
  - Training examples are relatively small.
  - Still a small number of mistakes or outliers exist.
  - Potentially important property pairs may not be characterized by training data.
- We use ensemble learning model
  - Base learner: decision tree
  - Random forest
    - Bagging
      - More suitable for relatively small training data
    - Randomness of input features
      - Potential important property pairs can be also used

Evaluation

- Integration in SView
  - SView finds candidates using owl:sameAs and web services
  - Light-weighted user interaction
    - Just accept or reject a candidate
  - Coreferent entities' data integrated in browsing immediately



Visit SView at http://ws.nju.edu.cn/sview/

Evaluation on SView dataset

- Target: evaluate how coCoref improves result quality using consensus partition
- Based on SView dataset (Oct. to Dec. 2013)
  - 36 users' individual partitions
  - 1,489 entities from 76 distributed data sources (URI namespace)
- Building reference partition
- Comparison using Precision, Recall, F-measure, Rand Index, NMI (normalized mutual information)
- Baseline: average measure values of 36 users' individual partitions
- Compare consensus partition with the best individual ones

Evaluation on SView dataset

- Results
  - Consensus partition is more comprehensive
    - Highest values for Recall, F-measure, Rand Index, NMI
  - Consensus partition is more robust
    - Improve low average accuracy of 36 users' results by 21% (0.807 to 0.665)
    - Achieve close precision to the best (0.807 vs 0.863)

	Precision	Recall	F-measure	Rand Index	NMI
Baseline	0.665	0.071	0.120	0.012	0.105
Best-of-K with highest F-measure	0.773	0.201	0.319	0.021	0.208
Best-of-K with highest Rand Index	**0.863**	0.186	0.306	0.090	0.439
Best-of-K with highest NMI	0.708	0.107	0.186	0.088	0.446
coCoref	0.807	**0.297**	**0.434**	**0.997**	**0.951**

Evaluation on OAEI NYT

- Target: evaluate the effectiveness of consensus partition and ensemble learning algorithm on a large scale
- Based on OAEI New York Times benchmark
- Leveraging coreference results of ObjectCoref, Zhishi.links, Knofuss
  - Total 33,914 entities
  - Cluster coreferent entities of each tool's dataset to construct it partition

Evaluation on OAEI NYT

- Results (1)
  - Highest accuracy of consensus partition

Table 2. F-measure comparison on the OAEI NYT test

	Consensus partition	ObjectCoref	Zhishi.links	Knofuss
NYT–DBpedia loc.	**0.948**	0.859	0.910	0.891
NYT–DBpedia org.	**0.939**	0.882	0.900	0.916
NYT–DBpedia peop.	**0.985**	0.958	0.970	0.960
NYT–Freebase loc.	**0.951**	0.938	0.882	0.913
NYT–Freebase org.	**0.959**	0.901	0.870	0.889
NYT–Freebase peop.	**0.988**	0.973	0.926	0.942
NYT–Geonames loc.	0.937	**0.938**	0.910	0.878

Evaluation on OAEI NYT

- Ensemble learning based on consensus partition
- Use 10% of training data
  - 10-fold cross validation
- Compare ensemble learning and consensus partition

Evaluation on OAEI NYT

- Results (2)
  - Identify a similar size of coreferent entities as consensus partition using only 10% of training data

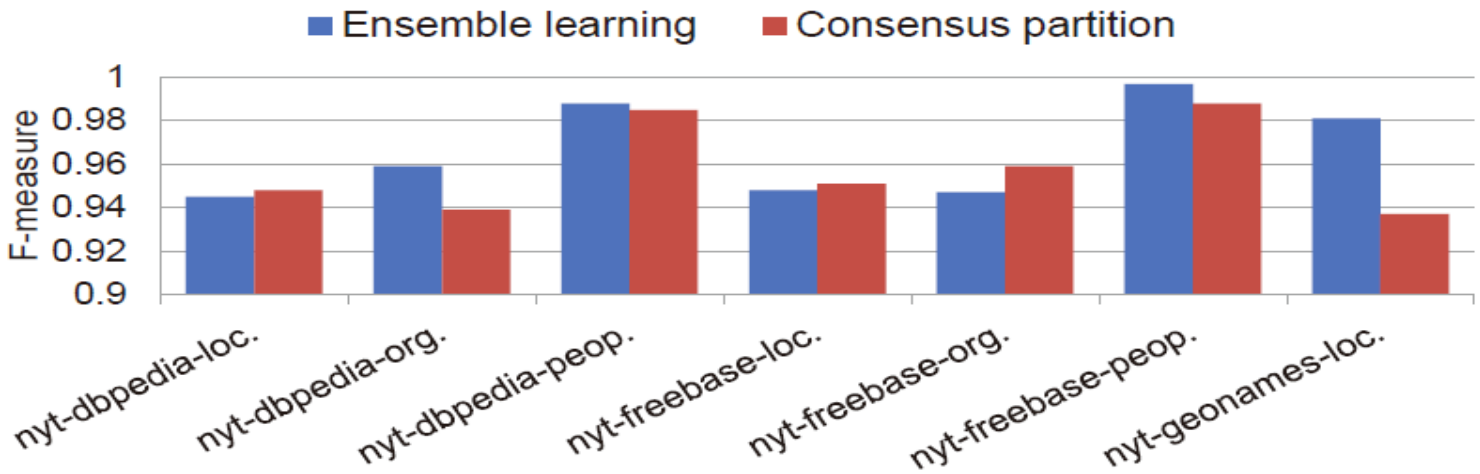


Fig. 3. Performance of ensemble learning

Conclusion

- Our contributions
  - Consensus partition for improving accuracy of user-judged results
  - Ensemble learning to alleviate user involvement
  - Integration with browsing activities

Thank you for your attention!

Supported in part by the National Natural Science Foundation of China under Grant Nos. 61370019 and 61170068, and in part by the Natural Science Foundation of Jiangsu Province under Grant Nos. BK2011189 and BK2012723