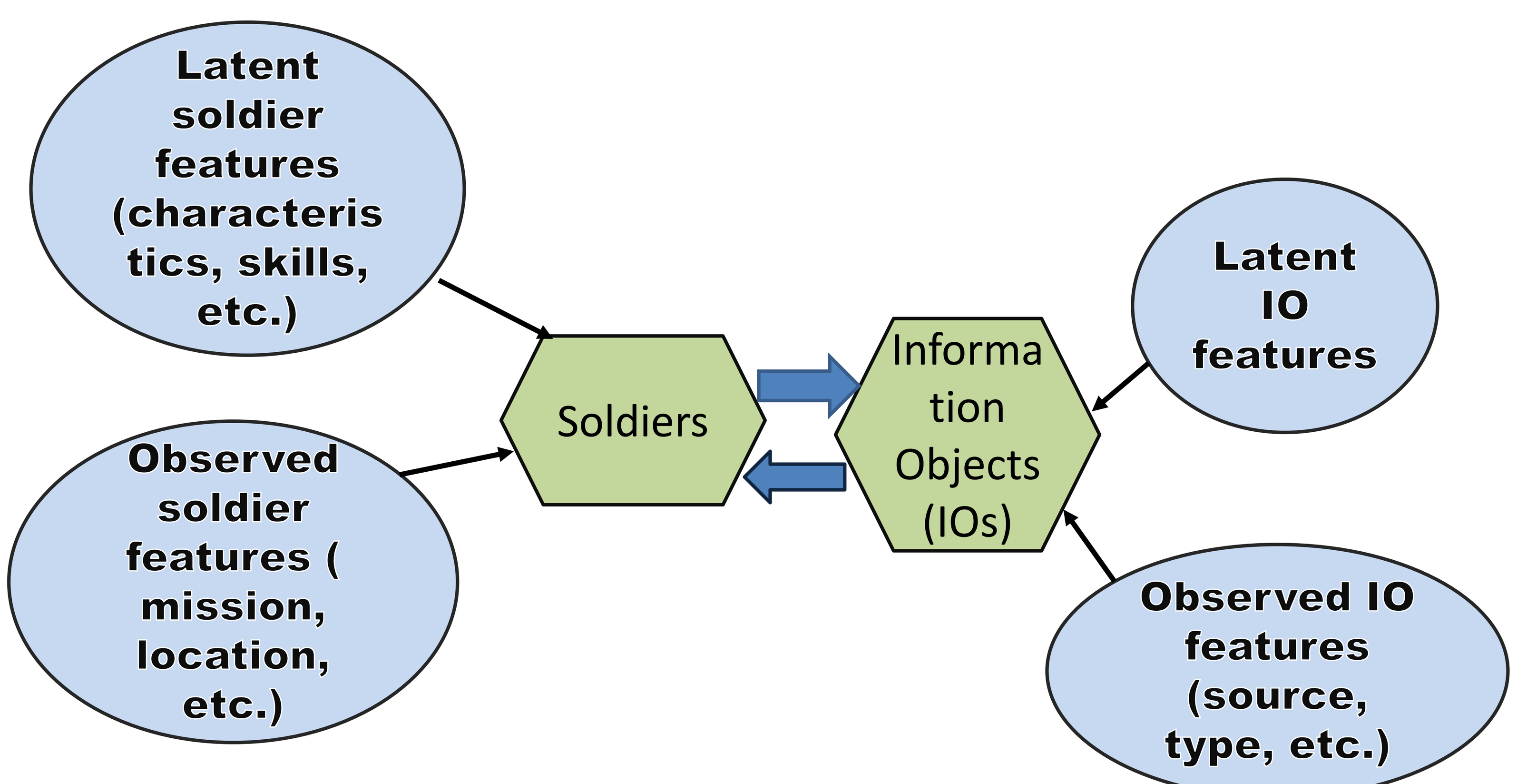
RESEARCH ON DETERMINING ADAPTIVE VALUE OF INFORMATION

ARL Mentor: Jade Freeman jade.l.freeman2.civ@mail.mil
Summer Intern: Konstantinos Pantazis kpantazi@umd.edu
Directorate: CISD

Research Objective

- Towards addressing the problem on information overload in tactical environment, this research seeks to find an algorithmic approach to quickly target critical operational information for dissemination over C3I networks.
- The **main objective** is to considering Value of Information (VoI) problem as a **recom-mendation task problem** and hence, incorporate state-of-the-art methods from the Recommender Systems' literature.



Soldier - IO feedback:

$$\text{Reward} = f(\text{Soldier}, IO) + \text{Noise}$$
where f similarity function.

Motivation

- Recent advancements in Data collection techniques have resulted to low-cost, large-scale and heterogeneous datasets to which tactical networks are unable to process in a timely manner.
- The notion of the ***value of information*** (VoI) is developed to support effective com-munications in tactical networks.
- VoI-based methods focus on the challenging problem of prioritizing and filtering valuable information to and from dis-mounted soldiers in a battlefield environment.

Background

- *Policy-based* VoI approaches have been developed to combat information overload and to reduce bandwidth based on *a priori* mission contexts using deterministic approach.
- Recently, automated procedures such as *Contextual Multi-Armed Bandits* (CMAB) have been proposed to complement existing approaches. Yet, they do not take into account
  - soldier relational data,
  - arrival of new soldiers and
  - time-evolving environments.

Proposed methodology

1. Develop a graph representation which encodes the mutual relations among soldiers [1] and implement **collaborative filtering** via matrix factorization techniques to account for the relations among soldiers into the reward function.

$$\text{Reward} = f(\text{Soldier's peers}, IO) + f(\text{Soldier}, IO) + \text{Noise}$$

2. Incorporation of a *drift term w* [3] to simulate changes in soldier's preferences over time.

$$\text{Reward} = w\Big[f(\text{Soldier's peers}, IO) + f(\text{Soldier}, IO)\Big] + \text{Noise}$$

3. Integration of Combinatorial CMAB schemes such as CUCB [2] for multiple-object recommendation.

Future work

- Appropriately integrate the proposed methodology into a *learning algorithm (Combinatorial CMAB)* for prioritizing and filtering IO in dynamic tactical networking environments.
- Provide desirable theoretical guarantees such as *optimal regret bounds.*
- Evaluate algorithm's performance empirically and/or theoretically based on simulations.

Acknowledgements

- I thank Dr. Jade Freeman for her guidance and help throughout summer.
- I also thank ARL's internship program for providing students the opportunity to participate in meaningful and hands-on research.

References

[1] H. Wang, Q. Wu, H. Wang. *Factorization Bandits for Interactive Recommendation*. AAAI 2017

[2] W. Chen, Y. Wang, Y. Yuan. *Combinatorial Multi-Armed Bandit: General Framework, Results and Applications.* PMLR 2013

[3] Q. Wang, C. Zeng, W. Zhou, T. Li, L. Shwartz, G. Y. Grabarnik. *Online Context-Aware Recommendation with Time Varying Multi-Armed Bandits.* KDD'16