Trust Amongst Rogues?

A Hypergraph Approach for Comparing Clandestine Trust Networks in MMOGs

Muhammad Aurangzeb Ahmad (Minnesota)
@bkeegan **Brian Keegan** (Northwestern)
Dmitri Williams (Southern California)
Jaideep Srivastava (Minnesota)
Noshir Contractor (Northwestern)

Context

Motivation

Immersion drives complex social behavior in MMOGs

Problematic uses and organizations in social media

see also: Lee, Eoff, Caverlee 2011

Gold farming

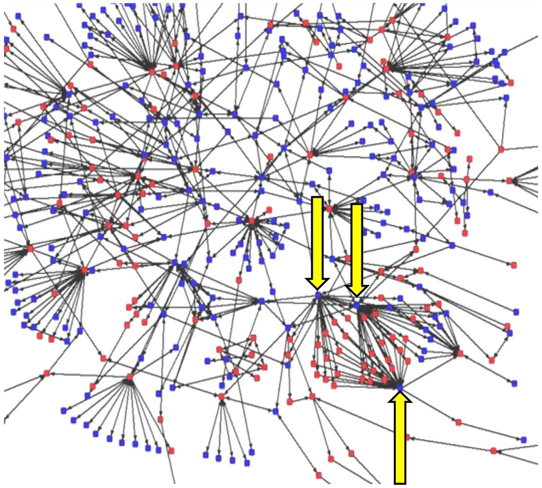
- ***Gold farming*** and ***real money trade*** involve the exchange of virtual in-game resources for "real world" money
- Laborers in China and S.E. Asia paid to perform repetitive in-game practices ("farming") to accumulate virtual wealth ("gold")
- Western players purchase farmed gold to obtain more powerful items/abilities and open new areas within the game
- Market for real money trade exceeds $3 billion annually **[Lehdonvirta & Ernkvist 2011]**

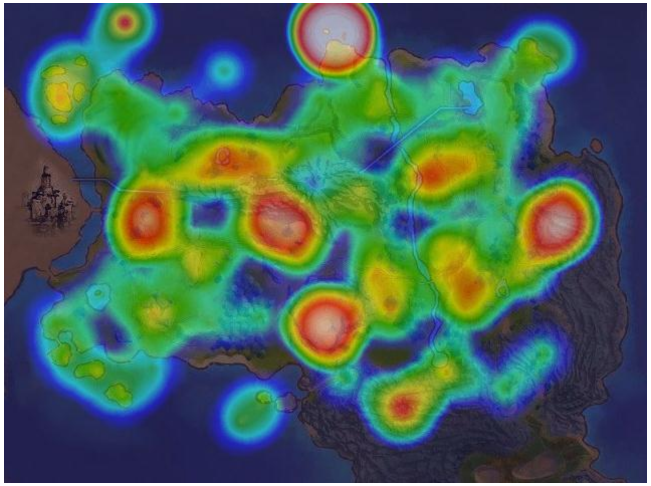
Game administrators ban gold farmers because of complicated implications

Clandestine behavior immaculately recorded in server logs

Structure of gold farmer trading networks similar to offline drug trafficking

[Keegan, Ahmad, et al. 2010, 2011]





High FPs in behavioral models → division of labor and undetected affiliates

[Ahmad, Keegan, et al. 2009]

Research Question

“A plague upon’t when thieves cannot be true one to another!”

– Sir Falstaff, *Henry IV, Part 1*, II.ii

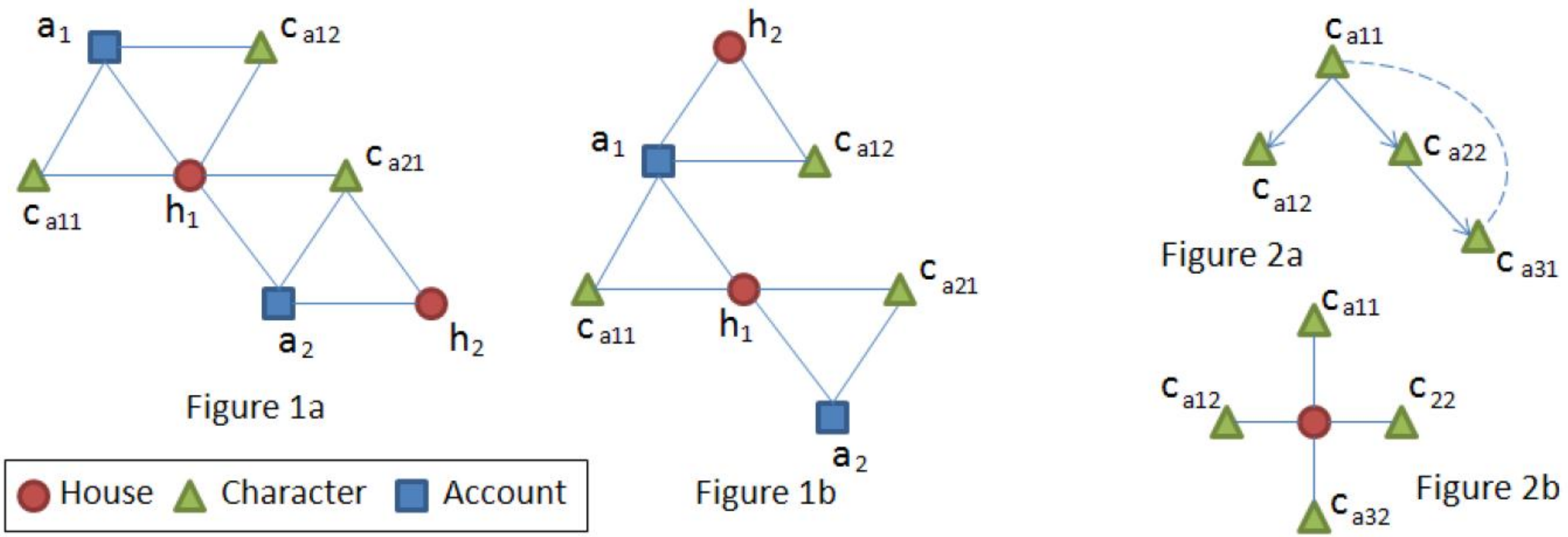
Do gold farmers trust each other?

Housing-Trust in EQ2

- Access permissions to in-game house as trust relationships
  - **None:** Cannot enter house.
  - **Visitor:** Can enter the house and can interact with objects in the house.
  - **Friend:** Visitor + move items
  - **Trustee:** Friend + remove items

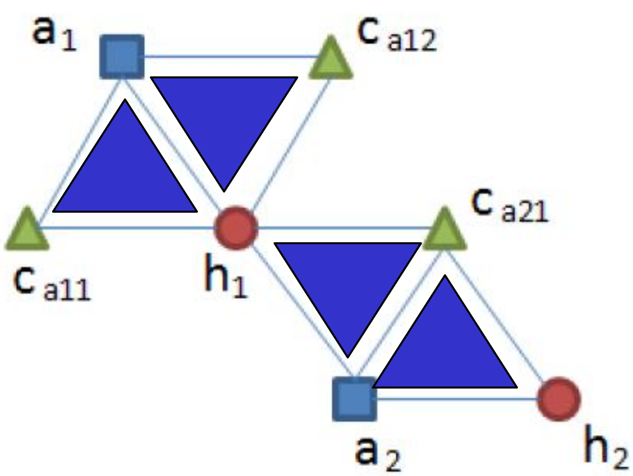
- Houses can contain also items which allow sales to other characters without exchanging on the market

Hypergraphs to Represent Tripartite Graphs



- Accounts can have several characters
- Houses can be accessed by several characters
- Projecting to one- or two-model data obscures crucial information about embededdness and paths
  - Figure 2a: Can c_{a31} access the same house as c_{a11}?
  - Figure 2b: Are characters all owned by same account?

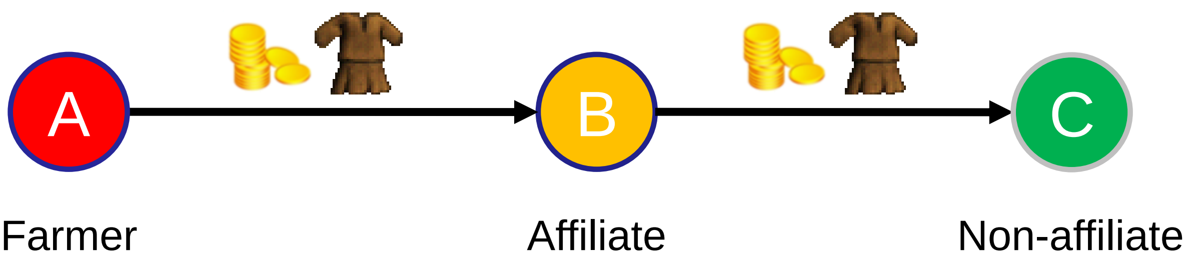
Hypergraphs: Key Concepts



- **Hyperedge:** An edge between three or more nodes in a graph. We use three *types* of nodes: Character, account and house
- **Node Degree:** The number of hyperedges which are connected to a node
  - $ND_{h1} = 3$
- **Edge Degree:** The number of hyperedges that an edge participates in
  - $ED_{a1\text{-}h1} = 2$

Approach

- Game administrators miss gold farmers and deviance is not a simple binary classification task
- **Guilt by association**: Identify “affiliates” who have ever interacted with identified gold farmers, but have **not** been identified as gold farmers themselves



EQ2 Dataset

- January 1, 2006 to August 31, 2006
- 38,217 characters
- 12,667 accounts
- 43,548 houses
- 3,013,741 hyperedges
- 151 accounts banned for gold farming (1.19%)

Network Characteristics

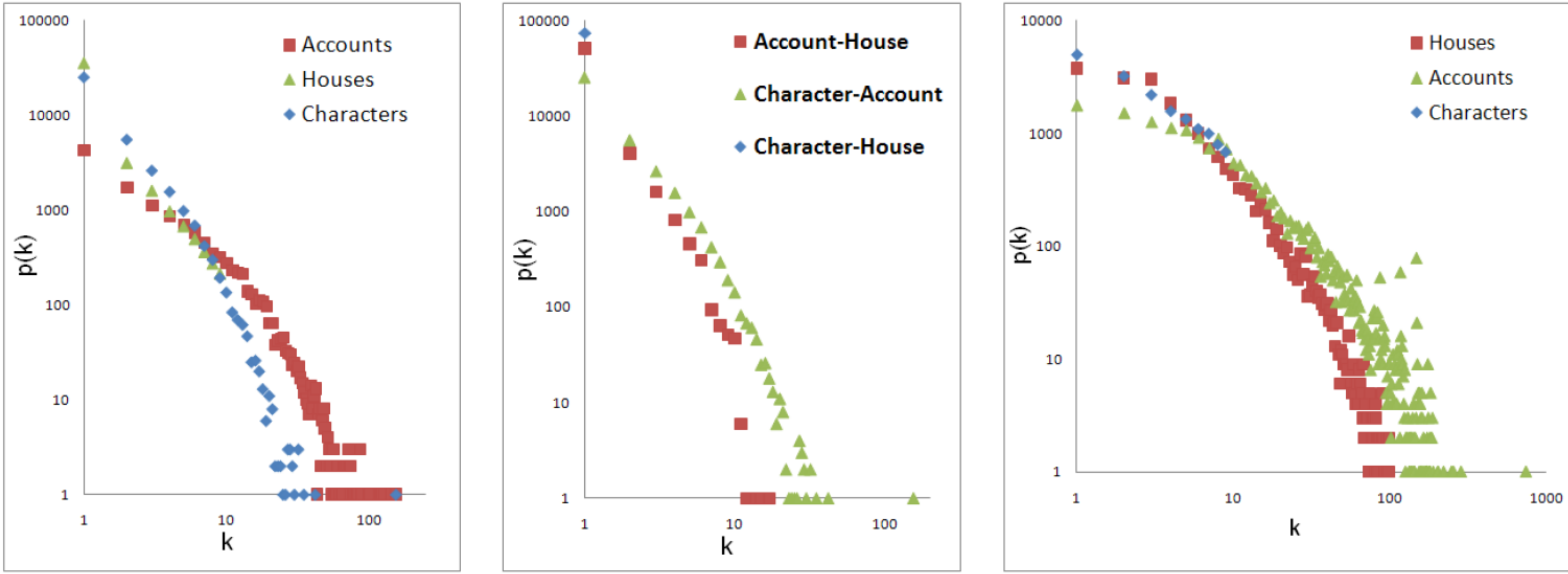


Figure 3a. Distribution of node degree for the trust hyper graph.

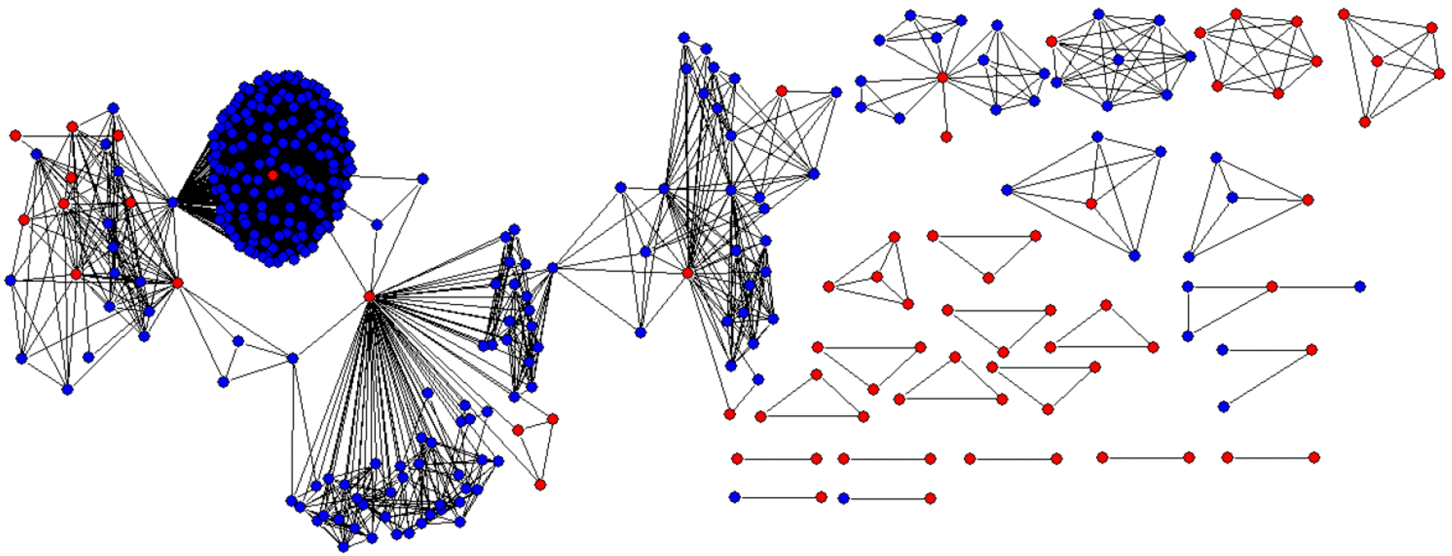
Figure 3b. Distribution of edge degree for the trust hypergraph

Figure 3c: Distribution of the projection networks

- Long tail distributions are observed for the various degree distributions
- The mapping from character-house to an account is always unique

Characteristics of Hypergraph Projection Networks

- **Account Projection:** Majority of the gold farmer nodes are isolates (79%). Affiliates well-connected (8.89) vs non-affiliates (3.47)
- **Character Projection:** Majority of the gold farmer nodes are isolates (84%). Affiliates well-connected (10.42) vs non-affiliates (3.23)
- **House Projection:** 521 gold farmer houses. Most are isolates (not shown) but others are part of complex structures. Densely connected network with gold farmers (7.56) and affiliates (84.02)



Key Observations

- **Picky picky:** Gold farmers grant trust ties less frequently than either affiliates or general players
- Gold farmers grant and receive fewer housing permissions (1.82) than their affiliates (4.03) or general player population (2.73)

	Total degree			In degree			Out degree		
	$< n >$	$< n_{GF} >$	$< n_{Aff} >$	$< n >$	$< n_{GF} >$	$< n_{Aff} >$	$< n >$	$< n_{GF} >$	$< n_{Aff} >$
Farmers	1.82	0.29	1.82	0.89	0.29	0.89	1.07	0.29	1.07
Affiliates	4.03	1.28	0.70	1.55	0.75	0.70	2.88	0.63	0.70
Non-Affiliates	2.73	-	7.77	1.57	-	5.98	1.56	-	2.34

Key Observations

- **No honor among thieves**
  - Gold farmers also have very low tendency to grant other gold farmers permission (0.29)
  - Affiliates also unlikely to trust other affiliates (0.70)

	Total degree			In degree			Out degree		
	$\langle n \rangle$	$\langle n_{GF} \rangle$	$\langle n_{Aff} \rangle$	$\langle n \rangle$	$\langle n_{GF} \rangle$	$\langle n_{Aff} \rangle$	$\langle n \rangle$	$\langle n_{GF} \rangle$	$\langle n_{Aff} \rangle$
Farmers	1.82	$\boxed{0.29}$	1.82	0.89	0.29	0.89	1.07	0.29	1.07
Affiliates	4.03	1.28	$\boxed{0.70}$	1.55	0.75	0.70	2.88	0.63	0.70
Non-Affiliates	2.73	-	7.77	1.57	-	5.98	1.56	-	2.34

Key Observations

- Affiliates are brokers:
  - Farmers trust affiliates more (1.82) than other farmers (0.29)
  - Affiliates trust farmers more (1.28) than other affiliates (0.70)
  - Non-affiliates have a greater tendency to grant permissions to non-affiliates (7.77) than in general (2.73)

	Total degree			In degree			Out degree		
	$< n >$	$< n_{GF} >$	$< n_{Aff} >$	$< n >$	$< n_{GF} >$	$< n_{Aff} >$	$< n >$	$< n_{GF} >$	$< n_{Aff} >$
Farmers	1.82	0.29	1.82	0.89	0.29	0.89	1.07	0.29	1.07
Affiliates	4.03	1.28	0.70	1.55	0.75	0.70	2.88	0.63	0.70
Non-Affiliates	2.73	-	7.77	1.57	-	5.98	1.56	-	2.34

Frequent Pattern Mining: Key Terms

Market Basket Transaction dataset example

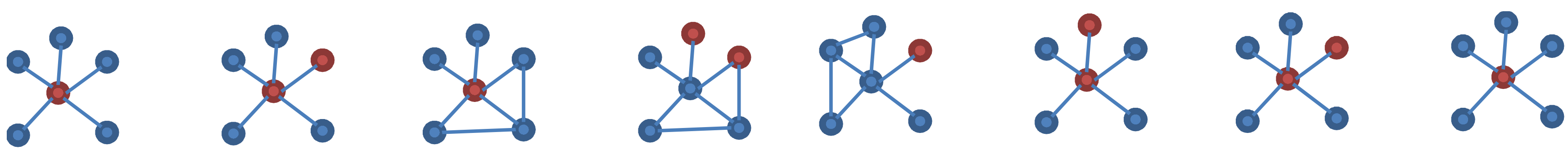
t1:	Beer, Diaper, Milk
t2:	Beer, Cheese
t3:	Cheese, Boots
t4:	Beer, Diaper, Cheese
t5:	Beer, Diaper, Clothes, Cheese, Milk
t6:	Diaper, Clothes, Milk
t7:	Diaper, Milk, Clothes

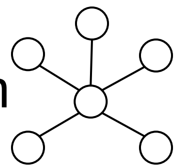
- **Items:** Cheese, Milk, Beer, Clothes, Diaper, Boots
- **Transactions:** $t_1, t_2, \ldots, t_n$
- **Itemset:** {Cheese, Milk, Butter}
- **Support of an itemset:** Percentage of transactions which contain that itemset
- Support(*{Diaper, Clothes, Milk}*) = 3/7

Frequent Itemset Mining for Frequent Hyper-subgraphs

- **Support of a Hyper-subgraph:** Given a sub-hypergraph of size *k*, sub_P is the pattern of interest containing the label *P*, sh_P is a pattern of the same size as sub_P and contains the label *P*, the support is defined as follows:

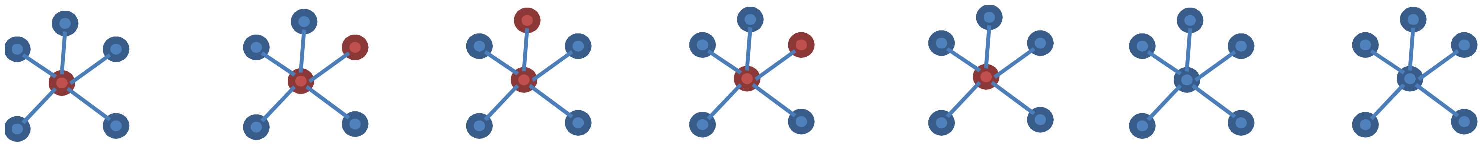
$$S = \frac{|sub_P|}{|\{sh_P \mid sh_P \subseteq H, |sh_P| = k\}|}$$

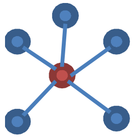


Support of pattern  also containing a gold farmer (red) = 5/8

- **Confidence of a Hyper-Subgraph:** Given a sub-hypergraph of size *k*, *subP* is the pattern of interest containing the label *P*, *subG* is a pattern which is structurally equivalent but which does not contain the label *P*, the confidence is defined as follows:

$$S = \frac{|sub_P|}{|\{sub_G | sub_G \subseteq H, |S| = k\}|}$$



Confidence of pattern  and containing a gold farmer = 5/7

Frequent Patterns of GFs

- Less than 0.1 support and confidence for almost all (except 8) frequent patterns with gold farmers
- Remaining 8 patterns can be used for discrimination between gold farmers and non-gold farmers
- Gold farmers & affiliates are more connected: A third of more complex patterns (*k* >= 10 nodes) are associated with affiliates (15/44)

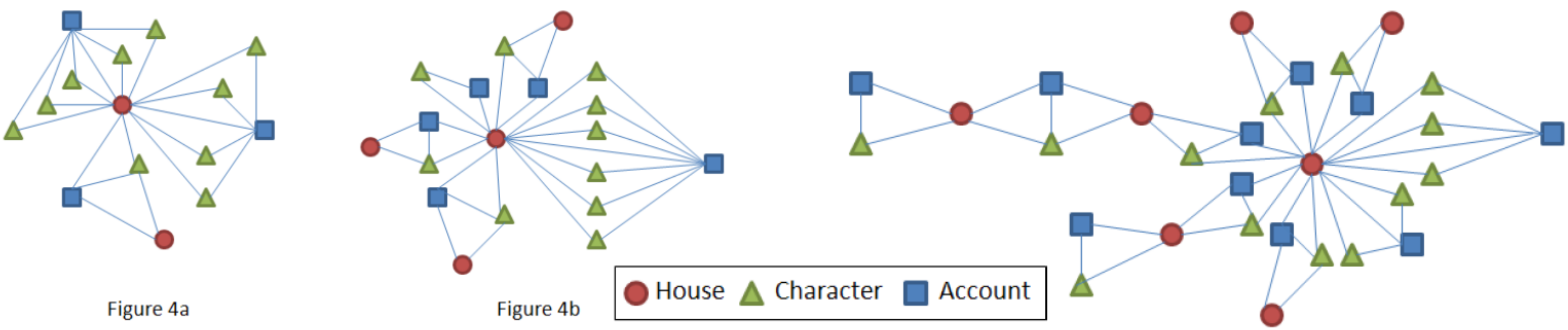
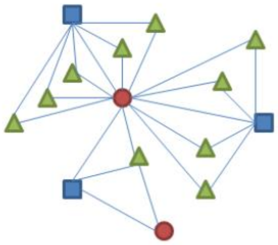
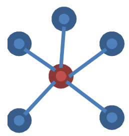


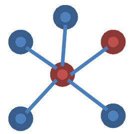
Figure 4a. Gold Farmer Hypergraph Pattern: Support = 0.33, Confidence = 1 Figure 4b. Gold Farmer Hypergraph Pattern: Support = 0.50, Confidence = 1 Figure 4c. Gold Farmer Affiliate Hypergraph Pattern: Support = 0.50, Confidence = 1

Conclusion and contributions

Using hypergraphs to represent complex data structures and dependencies



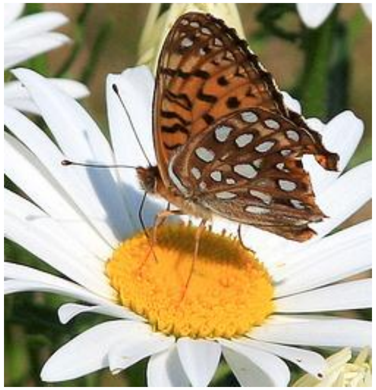




Application of frequent pattern mining to discover distinct trust patterns associated with gold farmers

No honor between thieves: Gold farmers tend not to trust other gold farmers

Implications



Social organization and behavioral patterns of clandestine activity as co-evolutionary outcomes

Using online behavioral patterns to inform and develop metrics/algorithms for detecting offline clandestine activity

Clandestine networks as “dual use” technologies – ethical and legal implications of improving detection?

[Keegan, Ahmad, et al. 2011]

Limitations and future work

- Housing/trust ties mediated by other or multiplex relationships
  - Communication, grouping, mentoring, trading, etc.
- Multiple types of deviance and deviants: Modeling role specialization & division of labor
- Using frequent subgraphs patterns as discriminating features for ML models
- Changes in frequent subgraphs over time
-

Thank you and questions

http://www.vwobservatory.org/

Gold Farming Related Publications

- Brian Keegan, Muhammad Aurangzeb Ahmad, Dmitri Williams, Jaideep Srivastava, Noshir Contractor. *Sic Transit Gloria Mundi Virtuali? Promise and Peril at the Intersection of Computational Social Science and Online Clandestine Organizations* The Third ACM WebSci Conference, Koblenz, Germany June 14-17, 2011 **(Best Paper Award)**
- Brian Keegan, Muhammad Aurangzeb Ahmad, Dmitri Williams, Jaideep Srivastava, Noshir Contractor (2011). *Using ERGMs to Map Online Clandestine Behavior to Offline Criminal Activity*. Sunbelt (XXXI) Florida February 8-13, 2011
- Brian Keegan, Muhammad Aurangzeb Ahmad, Dmitri Williams, Jaideep Srivastava, Noshir Contractor (2011). *Mapping Gold Farming Back to Offline Clandestine Organizations: Methodological, Theoretical, and Ethical Challenges*. Game Behind the Game. **(Best Paper Award)**
- Brian Keegan, Muhammad Aurangzeb Ahmad, Dmitri Williams, Jaideep Srivastava, Noshir Contractor, *Dark Gold: Statistical Properties of Clandestine Networks in Massively-Muliplayer Online Games* IEEE Social Computing Conference (SocialCom-10) Minneapolis, MN, USA, August 20-22, 2010.
- Muhammad Aurangzeb Ahmad, Brian Keegan, Jaideep Srivastava, Dmitri Williams, Noshir Contractor, *Mining for Gold Farmers: Automatic Detection of Deviant Players in MMOGS* Proceedings of the 2009 IEEE Social Computing (SocialCom-09). Symposium on Social Intelligence and Networking (SIN-09). Vancouver, Canada, August 29-31, 2009.

Frequent Pattern Mining: Key Termsm

- ***Goal:*** *Discover patterns which occur frequently in the data*
- *of* ***Set of items:*** $I=\{I_1, I_2, \ldots, I_m\}$
- ***Transactions:*** $D=\{t_1, t_2, \ldots, t_n\}$, $t_j \subseteq I$
- ***Itemset:*** $\{I_{i1}, I_{i2}, \ldots, I_{ik}\} \subseteq I$
- ***Support of an itemset:*** Percentage of transactions which contain that itemset.

Mining Frequent hyper-subgraphs in EQ2 Trust Data

- Since the data already contains a particular substructure *i.e.*, triads (house-account-character), this observation can be exploited for sub-hypergraph discovery
- We employ a “flattening” approach for representing the hypergraph

Mining Frequent hyper-subgraphs in EQ2 Trust Data

- ExpandG(h_set)

  For each house

  – Determine the accounts associated with it (if *d* <= *d_max*)
  – Get all the characters-account-house triples (c_i, a_i, h_i)
  – For each of the accounts
    - Determine the houses (*h_set_current*) associated with it
    - ExpandG(*h_set_current*)

- Lexicographically order all the triples associated with each house, a set of such triples becomes an individual transaction
- Apply standard association rule mining techniques