# Dark Gold: Statistical Properties of Clandestine Networks in Massively Muliplayer Online Games


Brian Keegan
School of Communication
Northwestern University
Evanston, IL 60201 USA
bkeegan@northwestern.edu

Muhammad Aurangzeb Ahmed
Dept. of Computer Science and Engineering
University of Minnesota
Minneapolis, Minnesota 55455 USA
mahmad@cs.umn.edu

Dmitri Williams
Annenberg School of Communication
University of Southern California
Los Angeles, CA 90089
dmitri.williams@usc.edu

Jaideep Srivastava
Dept. of Computer Science & Eng.
University of Minnesota
Minneapolis, Minnesota 55455 USA
srivasta@cs.umn.edu

Noshir Contractor
School of Communication
Northwestern University
Evanston, IL 60201 USA
nosh@northwestern.edu



***Abstract*—Gold farming is a set of illicit practices for gathering and distributing virtual goods in online games for real money. Using anonymized data from a popular online game to construct networks of characters involved in gold farming, we examine the trade networks of gold farmers, their trading affiliates, and uninvolved characters at large. Our analysis of these complex networks' connectivity, assortativity, and attack tolerance demonstrate farmers exhibit distinctive behavioral signatures which are masked by brokering affiliates. Our findings are compared against a real world drug trafficking network and suggest similarities in both organizations' network structures reflect similar effects of secrecy, resilience, and efficiency.**



***Keywords – dark networks, network analysis, online games, MMOG, MMORPG, EverQuest 2, gold farming, real money trade, cybercrime, deviance, scale-free, assortativity, attack tolerance***


## I. INTRODUCTION

Networks have assumed increasing importance as a both a theoretical and methodological approach toward understanding organizational structure and behavior [1, 2]. In particular, network analysis methods are potentially powerful tools to understand how actors in "dark" networks, such as drug traffickers and terrorist cells, coordinate their activities and adapt their structure to achieve their mission while avoiding detection and maintaining resilience [3, 4]. In practice, the hidden nature of the relationships in dark networks necessarily implies that collecting and analyzing complete or even representative data on these networks is very difficult.

However, the explosion of behavioral data available in online databases has opened up new avenues of social research [5]. One such source are massively multiplayer online games (MMOGs), large-scale social environments in which players of varying levels of expertise join cooperative teams to accomplish complex tasks [6, 7]. To the extent that individuals in online virtual worlds engage in similar psychological, social, and economic behavior as they do in the "real" world, virtual world research can potentially be mapped backwards and employed to understand real world phenomena [8]. Moreover, because the organizations that operate MMOGs maintain archival databases of all player actions and attributes, it is possible to analyze comprehensive cross-sectional and longitudinal behavioral data on a scale that would be unethical, impracticable, or impossible to do in the real world.

Using a combination of comprehensive, unobtrusively obtained data and methods in network analysis and data mining, we examine the coordination structures and dynamics of a dark network of one particular type of deviant activity in an MMOG. These "gold farming" networks operate under similar constraints as other criminal organizations, and so we argue that the structure and dynamics of these organizations can be used to characterize and understand deviant and criminal activity in other domains. We discuss our results and suggest implications for future work in organizational behavior, computational criminology, and intelligence.

## II. GOLD FARMING

### A. Background

Massively-multiplayer online games such as *World of Warcraft*, *EverQuest II*, and *Lord of the Rings Online* are examples of fantasy-based game worlds in which millions of players interact in a persistent virtual environment. While playing alone or with other players, they accumulate experience, armor, spells, and weapons to improve their power during encounters with non-player characters (NPCs) and player-versus-player combat (PvP). The virtual goods and in-game currency they acquire make their characters more powerful, and so their acquisition is typically one of the major goals of play.


The research reported herein was supported by the National Science Foundation via award number IIS-0729505, the Army Research Institute via award number W91WAW-08-C-0106, and the Air Force Research Laboratory by contract number FA8650-10-C-7010. The data used for this research was provided by Sony Online Entertainment.

However, the scarcity of these objects and the substantial investments of time needed to accumulate them drives demand high, and places a value on the time required to earn the goods and currency. Notably, many of these goods and currencies can also be obtained from other players within the game through trade and exchange. Unsurprisingly, just as these in-game economies exhibit similar macroeconomic characteristics observed in real-world economies [9], virtual worlds also contain black markets for acquiring goods and skills [10].

"Gold farming" and "real money trading" refer to practices that involve the sale of virtual in-game resources for real-world money via exchanges outside of the game itself. The name stems from a variety of repetitive routines ("farming") which are employed to accumulate virtual wealth ("gold") which is sold to other players who lack the time or desire to accumulate their own in-game capital [11, 12]. Gold buyers purchase this virtual capital to obtain more powerful weapons, armor, and abilities for their characters, accelerating them to higher levels and allowing them to explore larger parts of the game world and confront more interesting and challenging enemies, and increase their social standing [13].

Gold farming has been constructed as a deviant activity by both the game developers as well as the player communities for a variety of reasons. First, in-game economies are designed with carefully-calibrated activities and products that serve as sinks to remove money from circulation. Because gold farmers and buyers inject currency into the economy, they create inflationary pressure, unintended arbitrage opportunities, and other perverse incentives for market agents. Second, farmers' activities often overtly affect other players' experiences by excluding them from shared game environments, employing anti-social computer scripts ("bots") to automate the farming process, and engaging in the outright theft of account and financial information from their customers [11, 14]. Third, the game developers are risk-averse to the legal implications (such as property rights, taxation, and torts) of sanctioning a multinational industry estimated to generate between $100 million and $1 billion in revenue annually [15, 16] while lacking legal jurisdiction, precedent, or regulation [17, 18]. Finally, farming upsets the meritocratic and fantasy-based nature of the game in which some players can buy rather than earn accomplishments, thus potentially driving legitimate players away [19]. For these reasons, game developers actively and publicly ban accounts engaged in gold farming [20].

Although previous studies of gold farming, real money trade, and other forms of virtual property have examined player perceptions and proxy economic indicators [16, 21], like other illicit or illegal practices, organizational secrecy prevents the collection of reliable data. Moreover, rapid changes in practices and market demand, popular perceptions of gold farming as a frivolous novelty, significant language barriers, and geographic distance also prevent thorough observation or systematic examination [11].

### *B. Gold farming trade as a dark network*

Like other types of criminal organizations and dark networks, the pressure exerted by game developers to identify and disrupt gold farmers requires these clandestine operations to balance efficiency with security [22] and organizational resilience with operational flexibility [3]. While simple, routine, and unambiguous tasks are performed most efficiently in centralized network structures while difficult, complex, and ambiguous tasks are performed more efficiently in decentralized structures [23-25]. Thus, participants in dark networks attempting to maximize secrecy and efficiency must negotiate a dilemma in which decentralized networks provide the greatest security and resilience but low efficiency and flexibility [22].

Criminal networks' objectives and regularity of action likewise influence the structure they assume. Erickson's study of six diverse covert/secret organizations emphasizes that organizations with an established reputation are committed to emphasizing security over efficiency [23]. Baker and Faulkner's study of price fixing and collusion in the heavy electrical equipment industry revealed that peripheral players were less targeted and less sanctioned than more central players [24]. Decentralization has also been observed to be a key tactic adopted by members of a criminal network in response to targeting and asset seizure by law-enforcement [25]. Compared to the networks of al-Qaida terrorists, a drug trafficking enterprise engaging in regular activity exhibits higher centralization and a core of closely-linked participants with stable roles [22]. The addition of other actors to the core of a criminal network can serve to extend its periphery and insulate participants at the core [26].

If dark networks are organized to maximize concealment and secrecy, networks of gold farmers and their affiliates should have (1) lower centrality relative to non-farmers, (2) decentralized topologies, and (3) dissortative mixing to reduce the likelihood of identification and expulsion. Furthermore, the farming network should also exhibit substantial resilience to the removal of gold farmers from the network as a result of random failures as well as concerted attacks. These concepts are described and operationalized in greater detail below.

## III. DATA AND METHODS

### *A. Data and preparation*

Anonymized database dumps were collected from Sony Online Entertainment's massively-multiplayer online game *EverQuest II*. These data include both cross-sectional attribute data about individual characters as well as longitudinal data cataloging character-to-character transactions. Because activity within the game is spread amongst several unique servers running instances of the game in parallel, a record of player-to-player exchanges on a single, representative server was condensed to generate a weighted, directed edge list of all transactions between characters on that server between for 36 weeks between January and September 2006. A separate table recording instances of accounts banned by the developer for abuse, non-payment, and other reasons was parsed to extract rationales related to "plat", "spam", "farm", "gold", "coin", "bot", and "launder". The list of banned gold farmers was intersected with the trade activity database based on common account identification numbers to identify counterparty characters in which at least either the transaction sender or receiver had been banned by game administrators using gold farming-related rationales.

Just as a list of judgments from criminal proceedings is not an exhaustive account of all criminal activity, the cancellation table is not a complete list of all gold farmers. Previous research using a machine learning approach to classify gold farmers based upon demographic and behavioral variables generated a large number of false positives, which may be evidence of unidentified gold farmers in the data [27]. Given the presence of identified gold farmers, unidentified gold farmers, and non-gold farmers, we identify three distinct types of networks. The *farming network* is the set of all characters whose accounts have been identified as gold farmers by the administrators at any point in time and the trade relationships among them. The *affiliate network* is the set of all characters that have ever engaged in a trade transaction with a farming character as well as the farmers themselves and the relationships among both sets. Reasonable suspicion exists that the counterparties to gold farmers' trade interactions have a higher likelihood of being unidentified gold farmers. Finally, the *non-affiliate network* is the set of characters that have never interacted with identified gold farmers. All three networks are directed, which means that a tie from actor A to actor B is distinct from a tie from actor B to actor A.

We also use network data on a drug trafficking ring obtained from a Canadian law enforcement taskforce called Project Caviar to contextualize our findings for the gold farming network [25, 28]. Although orders of magnitude smaller in network size ($N$=110, $E$=295), the Caviar data is also directed longitudinal data of criminals and co-offenders which makes it a direct real-world analogue against which we can compare network statistics and dynamics.

### B. Network statistics: centrality, weight, clustering

Network analysis offers several statistical metrics for calculating the most "important" or "prominent" node in a network as a function of its connectivity profile. Specific definitions for each can be found in [29].

- *In-degree* and *out-degree centrality* reflect the number of incoming and out-going directed ties for a given node. Nodes with high in-degree centrality have many incoming links.
- *Closeness centrality* measures how close a given node is to the rest of the network, or the inverse of the sum of shortest paths to every other actor in the network. Nodes with high closeness centrality can reach every other node in the network in relatively few steps.
- *Betweenness centrality* measures the extent to which an actor lies on many shortest paths between every other node, connecting nodes that would not otherwise be connected through short or direct paths.
- *Eigenvector centrality* is a measure of prestige recursively calculated by taking a given node's influence as a function of the influence of the nodes connected to it. Nodes with high eigenvector centrality are themselves connected to other nodes with high eigenvector centrality.
- *Clustering coefficient* is estimated by measuring the extent to which a given node's counterparties have each other also as counterparties. A node with a high clustering coefficient in a trading network implies that many of its partners also trade with each other.

### C. Network dynamics: connectivity, assortativity, tolerance

Many social, collaboration, and technological networks exhibit highly centralized network topologies, principally characterized by a frequency distribution of individual nodes' degrees following a power law $P(k) \sim k^{-\gamma}$. Complex networks exhibiting this scale-free property are not generated randomly but emerge as a result of growth and preferential attachment of new nodes to existing nodes of high degree [30, 31]. In some centralized networks, the distribution is not a true power law. Truncated power laws have attenuated distriubtions in the tail which suggests that extrinsic factors are limiting the ability of high degree nodes to scale up efficiently. These effects may be related to aging (some old nodes stop receiving new links after some time threshold) or cost effects (maintenance of links has non-trivial marginal costs) [32].

Scale-free and other complex networks can also be characterized by the extent to which the degree of individual nodes is correlated with the degrees of its neighbors. Networks exhibiting *assortative mixing* are defined by high-degree actors being connected to other high-degree actors while low-degree actors are connected to other low-degree actors. Conversely, networks exhibiting *dissortative mixing* have high-degree actors connected to low-degree actors and low-degree actors connected to high degree actors. Newman defines a connected degree-degree Pearson correlation coefficient:

$$r = \frac{1}{\sigma_q^2} \sum_{jk} jk(e_{jk} - q_j q_k)$$

where $\sigma_q^2$ is the variance of the normalized distribution $q_k$ of neighbors' degrees for a given degree $k$. Social and collaborative networks tend to exhibit strong assortativity while technological and biological networks generally exhibit dissortativity [33, 34]. We expect that clandestine networks will exhibit dissortativity as highly-central actors will be more-strongly motivated to distance themselves from other highly-central actors so to maintain secrecy and resilience rather than preferentially joining groups and teams with other highly-central actors [35].

Finally, we examine the error and attack tolerance of the farmer and affiliate networks. Previous research has examined the extent to which scale-free and other complex networks remain robust even at unrealistically high failure rates [36, 37]. We similarly define *error tolerance* to be the extent to which the trade network remains connected despite random removal of nodes. Given the high variability of centralization in the network and prevalence of actors with low degree, it is improbable that any a random failure will fragment the network into disconnected subcomponents or isolated nodes. Thus, the random attack serves as a baseline against which we can assess the relative performance of the other attack strategies.

Because game companies actively attempt to identify and ban gold farming accounts from the game, we simulate two possible strategies for fragmenting the farming and affiliate networks to assess how quickly this network can be broken. Attacks employing *node degree targeting* sequentially remove

nodes with the highest connectivity (degree) in the network. Attacks employing *edge weight targeting* sequentially remove the dyads sharing the highest weighted edges on the network. Scale-free networks with heterogeneous distributions of connectivity or edge weight imply that connectivity across the network is maintained by a few highly connected nodes whose removal can drastically alter the network's topology [36]. The resilience of the network to attacks removing nodes can be assessed by the extent to which removal fragments the network. Fragmented networks can be characterized by two properties: having many unconnected nodes (isolates) and the largest subcomponent of the network having a relatively small fraction of the total nodes in the networks. We expect that gold farming network will exhibit substantial tolerance to random error as well as targeted attacks.

## IV. Results

### A. *Differences in centrality, weight, clustering*

The union of all three networks consists of 228,365 directed edges between 43,021 characters. The farming network has 1,604 nodes, 2,930 edges, and the maximum observed transactions between characters was 1,459. The affiliate network had 5,367 nodes, 29,178 edges, and the maximum flux remained 1,459.

Multinomial logit regression models for the three node types (farmer, affiliate, non-affiliate) were run to measure the differences between farmers' and affiliates' centrality scores against non-affiliates' centrality. Farmers had significantly lower in-degrees (β=-0.0473, z=-9.24) and out-degrees (β=-0.189, z=-5.11) while affiliates had significantly higher in-degrees (β=0.0626, z=47.21) and out-degrees (β=0.0529, z=44.25) as compared with non-affiliates respective degrees. Farmers and affiliates both also received ($\beta_{farmer}$=0.00691, z=23.04; $\beta_{affiliate}$=0.00851, z=32.64) and initiated ($\beta_{farmer}$=0.00796, z=24.32; $\beta_{affiliate}$=0.009556, z=32.97) more transactions than lay-characters. These significant differences corroborate our hypotheses and suggest farmers – rather than camouflaging their activity or counterparties by emulating typical players' trading practices – rely upon many repeated transactions with only trusted (and likely co-offending) characters. This behavior is contrasted by gold farmers' affiliates who have substantially higher connectivity and transaction frequency than lay players.

Farmers (β=-0.679, z=-6.38) and affiliates (β=-2.438, z=-13.24) had significantly lower closeness centrality scores than the unaffiliated population at large. This suggests farmers and their affiliates generally position themselves on the periphery of the transaction network such that most actors would have to pass through several intermediary characters to trade with them. Farmers' (β=1.56x10$^{-7}$, z=1.94) have only marginally greater betweenness centralities than the population at large, although while affiliates' betweenness (β=1.36x10$^{-6}$, z=35.81) is significantly greater. Farmers also had significantly lower eigenvector centrality (β=-10300, z=-8.60) while affiliates had significantly higher (β=5377, z=35.73) centrality as compared with non-affiliated characters. Affiliates are more likely to connect characters that cannot otherwise connect to each other while also being connected to by influential characters, demonstrating they serve a key brokering role effectively insulating characters with a higher likelihood of being banned from the game from characters in the rest of the network.

Finally, farming networks exhibit significantly higher clustering coefficients (β=0.905, z=8.64) than the non-affiliate population while affiliates do not have significantly different clustering coefficients (β=-0.0926, z=-0.81). Since farmers' trade networks are substantially more tightly-knit, the fact that counterparties are more likely to trade with each other provides a greater level of redundancy if any single node is removed. Affiliates' trading partners, on the other hand, are no more tightly-knit than the general population of characters. This is further evidence that although affiliates serve a unique role brokering connections between the farmers and typical characters, they exhibit structural characteristics of both farmers and non-affiliates.

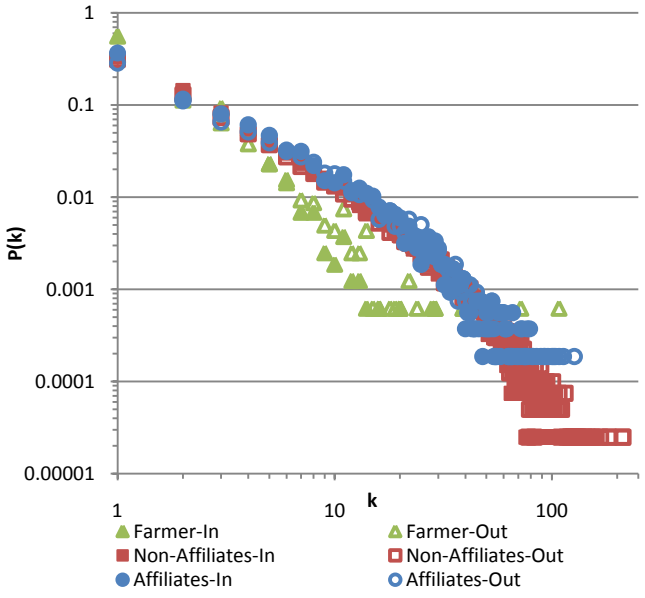


***Figure 1:*** In- and Out-degree distributions of trading networks

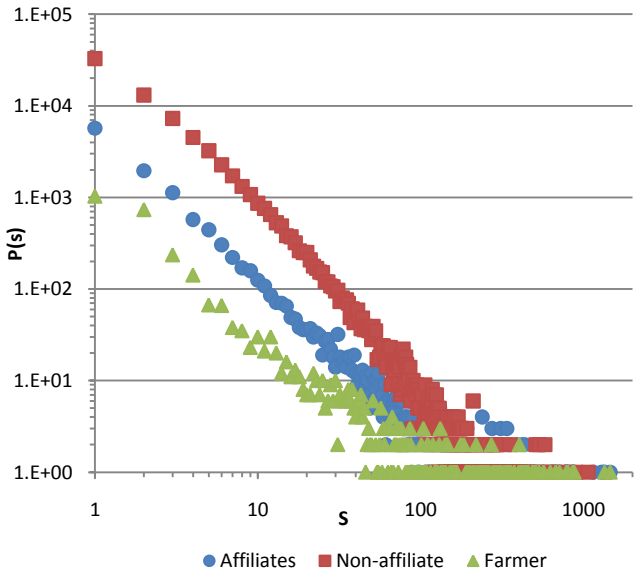


***Figure 2:*** Tie strength distribution of networks

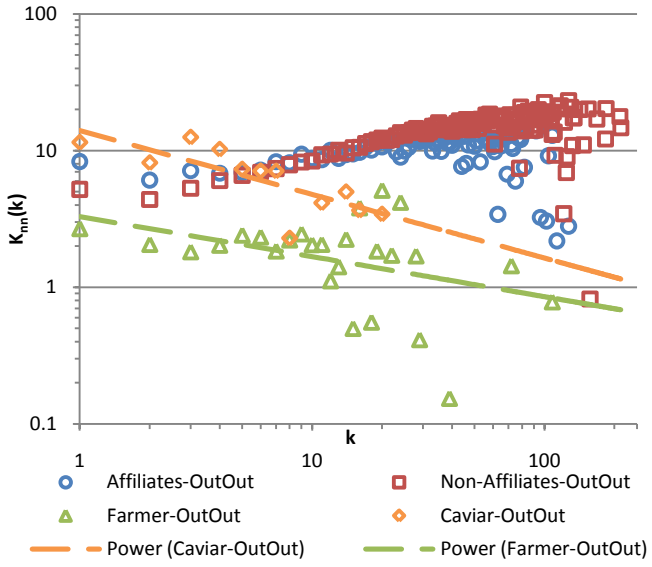


***Figure 3:*** Average neighbor out-degree for affiliate, non-affiliate, farmer, and Caviar networks against actor degree

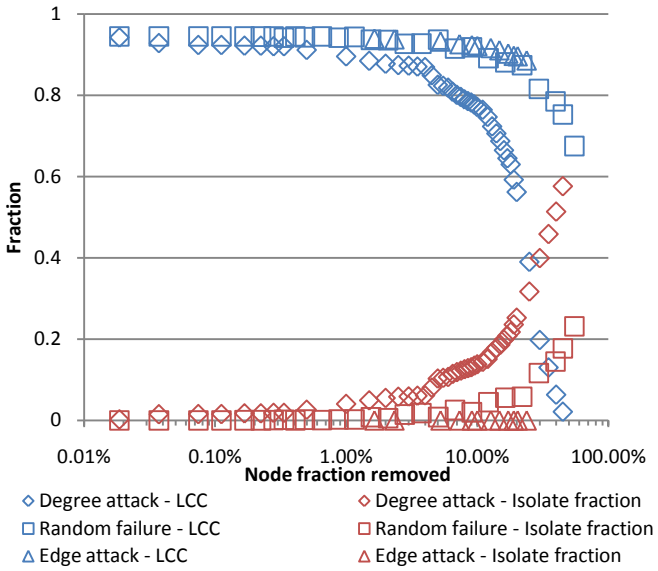


***Figure 4:*** Fraction of isolate actors, in red, and fraction of network actors in largest connected component (LCC), in blue, after node removal by random, degree, or edge attacks

### *B. Degree and weight distribution*

Figure 1 plots the in-degree and out-degree distributions of the farmer, affiliate, and non-affiliate networks. The quasi-linear relationship on log-log scale implies a scale-free structure of this network. These networks' heterogeneous connectivity implies a highly centralized network in which the majority of characters have few trading links with each other, but a handful of characters have several hundred or over a thousand counterparties. Although networks with scale-free degree distributions are generated by an underlying fitness function that causes new actors to preferentially link with existing actors, the degree and weight distributions observed in Figure 1 do not exhibit ideal scale-free behavior. In the degree distribution, we observe an attenuation of connectivity among higher-degree nodes in all three network types rather than a straight linear decrease throughout the domain. The onset of this truncated connectivity for non-affiliated and affiliated nodes occurs well above ($k \approx 20$) the average degree in the network ($k = 5.4$) while the fall-off occurs more rapidly for the farmer network ($k \approx 5$). Models of node aging and limited interaction capacity both approximate this truncation of scale-free connectivity: farmers may avoid over-exposure by limiting the number of trade relationships they have or preferentially linking to only those actors possessing specific attributes [32, 38].

Figure 2 plots the distribution of tie strengths for each of the three networks. Again, a power law distribution is tie strengths suggests that the vast majority of dyads in each network only transact once or twice, but there are several dozen dyads trading with each other more than 100 times over nine months. The intensification of tie strength is also observed in the farming network around a tipping point ($S \approx 25$). The shift above this threshold suggests there are more gold farmer dyads engaging in more intense trading than would be expected by extrapolation. This effect may be an adaptation that promotes operational efficiency, security, or resilience, thus high-frequency trading may potentially be a reliable behavioral signature for identifying gold farming transactions; these dyads will be the first removed in attack tolerance simulations below.

### *C. Assortativity*

Figure 3 plots the average degree ($k_{nn}$) for the nearest neighbors of nodes with a given degree $k$. The non-affiliate network is clearly assortative based on the positive relationship ($r$=0.201) between node degree and neighbor's degree. Frequent traders are more likely to interact with other frequent traders supporting our assumption that trade interactions are fundamentally social and collaborative. As hypothesized, we observe the subset of the affiliate network consisting of only identified farmers exhibits a clear pattern of dissortative mixing ($r$=-0.261) that resembles the dissortative pattern observed in the offline Caviar drug trafficking network ($r$=-0.466). The presence of dissortative mixing in both the drug trafficking and gold farming networks is key evidence in validating our assumptions that behaviors in online, virtual worlds also map onto behaviors found in the offline, real world.

Interestingly, the affiliate network exhibits characteristics of both the associative non-affiliate network and the dissortative clandestine networks ($r$=0.015). In Figure 3, for $1<k<30$, the network is clearly assortative, nearly matching the non-affiliate network. However, for $k>30$ the network exhibits substantially more variance including a cluster of high-degree outliers with significantly lower neighbor degree centralities. These findings suggest that the affiliate network exhibits characteristics of both the legitimate, non-affiliate trade network in the $k<30$ domain as well as the farmer trade network in the $k>30$ domain. These high degree nodes with low degree neighbors in the affiliate network likely include unidentified gold farming characters. These heterogeneous mixing distributions for the affiliate and farmer networks further motivate our analysis below to understand whether the removal of high-degree brokering characters can effectively fragment the network.

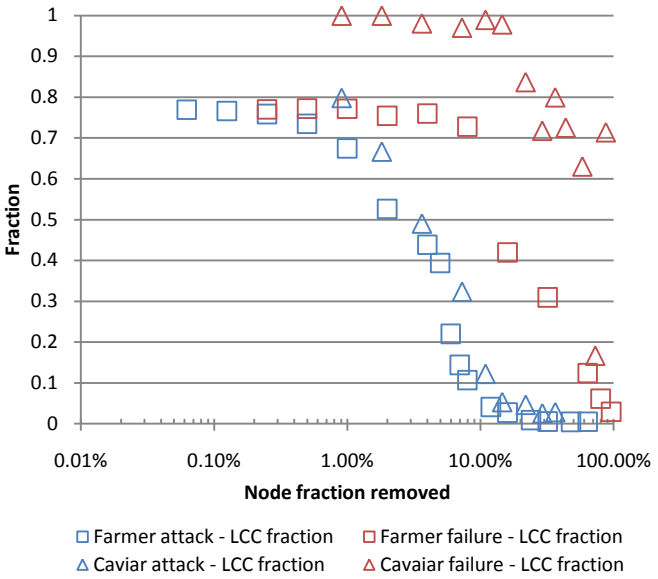


***Figure 5:*** Fraction of actors in largest connected component (LCC) of Farmer and Caviar networks

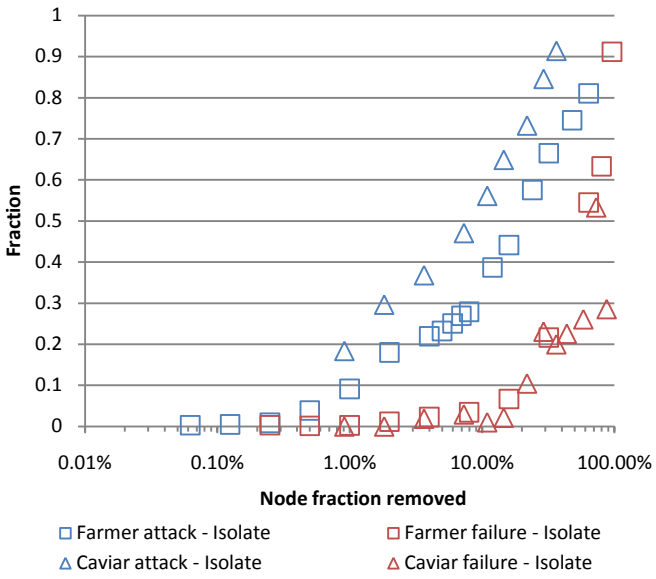


***Figure 6***: Fraction of isolated actors in network of Farmer and Caviar networks

### *D. Attack tolerance*

Attacks on the gold farming and affiliate networks were simulated using the random failure, degree attack, and edge attack targeting strategies. Figure 4 displays the results of these attacks on both the fraction of the affiliate network located inside the largest connected component (LCC) as well as the fraction of the affiliate network that are isolates. Notably, the edge attack strategy of extracting nodes with strong link weights performs more poorly than both degree attack and random node failure. When all nodes with edges strength greater than 16 transactions have been removed, the resulting network has 50% fewer links and 24% fewer actors, yet there were still no isolates in the network and the largest connected component still contained over 88% of network actors. The resilience of the network under an attack removing nodes based on their rank-ordered maximum tie strength – even at similar fractions of removed nodes that lead to fracturing under random and degree attacks – suggests the prevalence and distribution of weak ties in this affiliate network gives it substantial resilience.

The affiliate network is likewise resilient to degree attack as it only begins to appreciably fracture after removing 10-15% of the most well-connected characters. More than 30% of the most well-connected nodes must be removed to ensure that 50% or more of the network are isolates or not in the LCC. Because the affiliate network exhibits hybrid properties which makes it difficult to accurately and precisely distinguish gold farmers from legitimate characters, if even a fraction of the highest-degree affiliates are legitimate characters who cannot be justifiably removed, these genuine actors' residual connectivity is problematic. Administrators would need to identify then remove substantial fractions of the affiliate network containing currently unidentified farmers to effectively decompose it.

Figures 5 and 6 compare the identified farmer network to the Caviar drug-trafficking network. Both the real world and virtual criminal networks exhibit very similar performance and resilience under degree attack and random failures. Removing fewer than 1% of the nodes by attack keeps the fraction of the network in the LCC relatively high and the number of isolates in the network relatively low. However, these networks are an order of magnitude more sensitive to node removal than the affiliate networks analyzed in Figure 4; removing approximately 5% of nodes by degree attack cuts the fraction of nodes in the largest connected component below 50% while increasing the fraction of isolates to approximately 50%.

Taken together, this analysis shows the farmer and affiliate networks have substantial resilience to both random failures and determined attacks over several orders of magnitude before fracturing into many disconnected components, a pattern which is also found in a real-world drug trafficking network. The affiliate network composed of farmers, unidentified farmers, and legitimate players exhibits even less sensitivity to attack than the clandestine networks alone. These findings suggest that farmers are able to effectively conceal their interaction patterns against the background of legitimate trade activity which also provides substantial resilience to interdiction.

## V. DISCUSSION

### *A. Conclusions*

Gold farmers ply their trade on the periphery of a complex and heterogeneous trade network. Rather than interacting directly with the general population, farmers broker their transactions through a complex network of undetected affiliate characters. The significant differences in connectivity and assortativity between farmers, affiliates, and typical players suggests characters engaged in activities with a higher likelihood of being banned adapt their behavior and interactions to support the twin imperatives of efficiency and secrecy. Although farmers form fewer connections than lay characters, they trade very intensely within their highly-clustered immediate networks. The heterogenous connectivity of this network combined with its dissortative mixing appear to exhibit

a similar level of resilience to simulated attacks as observed in a real world drug trafficking networks. Gold farmers' trading counterparties likewise exhibit a centralized network with skewed connectivity and transaction frequency, but they also trade with many more characters than expected in the general character population. These affiliated characters appear to fulfill a crucial role in brokering trades between farmers and non-farmers and exhibit connective and assortative characteristics of both groups. However, this affiliate network requires a significant fraction of nodes to be removed before it fractures and the likely presence of unidentified farmers embedded amongst legitimate players further heightens the difficulty for administrators.

However, from a more theoretical perspective, the dynamic processes that generate both networks' characteristic structural patterns (e.g., heightened clustering, truncated degree distriubtions, dissortative mixing) imply that clandestine and deviant organizations in online settings are motivated and constrained by the similar organizational pressures observed in offline criminal organizations. Whereas prior attempts to employ social network analysis to understand criminal, deviant, or other clandestine networks have had difficulty overcoming the substantial barriers towards collecting reliable data on these organizations [39, 40], by observing patterns of relational data in a massively multiplayer online game, we were able to characterize complex properties of several thousand individuals in a clandestine networks.

However, these conclusions carry some caveats. While we assume the organizational behavior of gold farmers will reflect the similar imperatives and behavior as real-world criminal organizations like drug traffickers, gold farming networks also potentially differ from real world criminal networks because of the intrinsic affordances of electronic medium in which they operate. While traditional criminal organizations rely on trust to recruit co-offenders and engage in their tasks [41], gold farmers can effectively create new or replacement co-conspirators at low costs by registering new accounts. Thus, gold farmers, unlike other criminals, do not need to recruit or convert existing players, engage in processes of ideological identification or control over existing members, or threaten violence against defectors to accomplish their mission [3, 40].

The present analysis likewise offers only a cross-sectional and macro-level analysis of the dynamics and structures involved in clandestine organization. Our sampling of the data and subsequent analysis is also necessarily biased by the heuristics employed by the game developer to identify deviant players and almost certainly omits the interactions of some actors engaged in unidentified deviant acts. Our analysis likewise does not control for endogenous variables such as changes in features of the game play or player population which may give rise to significant changes in patterns and structures of interaction among players.

### *B. Implications and future work*

Our research demonstrates that actors engaged in deviant behavior in an online game operate under similar constraints and motivations in response to authorities' enforcement activities as offline criminal organizations. This is evidence that there is indeed some "mapping" between virtual and real criminal networks. However, given the potential of mining structural and behavioral data from large corpuses, characterization and analysis of criminal networks must go beyond cross-sectional and macro-level descriptive statistics of structure and address both the underlying dynamics of local-level structural change in response to enforcement actions. Other types of relational data present within the game such as communication patterns, group interactions, behavioral patterns, and trust proxies can be incorporated into developing more thorough analyses of multiplex relations among members of deviant organizations. Future research should develop machine learning models of structural and behavioral signatures likely to predict gold farming behavior. Furthermore, systematic disparities between rule sets predicting farmers and the from these models could then be used to identify biases in authorities' heuristics for identifying farmers as well as emergent properties of actors employing traits to exploit these gaps.

In particular, coevolutionary processes of variation, selection, and retention of behavioral signatures within these criminal organizations need to be reconciled with existing theories of criminal behavior [42, 43]. Given the variation in missions, members, and media across different criminal organizations, we suggest future research should incorporate meta-analyses of motifs and other structural features common to real world criminal organizations and networks. For example, in what respects are drug trafficking operations, terrorist cells, and white collar conspiracies similar or different with regard to variation and retention of structural and behavioral features under selection pressures? P*\exponential random graph models (ERGM) are potentially powerful tools for comparing disparate networks [44] and may provide insights into the extent to which the observed gold farming networks exhibit similar structural tendencies and can potentially can support the development of multilevel, multitheoretical models of criminal networks [45].

## VI. ACKNOWLEDGEMENTS

We thank Sony Online Entertainment for the EverQuest II data, Dr. Carlo Morselli for his Project Caviar data, members of the Virtual Worlds Observatory team for their feedback, and the National Science Foundation, Army Research Institute, and Air Force Research Laboratory for their support. The findings presented do not in any way represent, either directly or through implication, the policies of these organizations.

## VII. REFERENCES

[1] D. Brass, J. Galaskiewicz, H. Greve, and W. Tsai, "Taking stock of networks and organizations: A multilevel perspective," *Academy of Management Journal*, vol. 47, no. 6, 2004, pp. 795-817.

[2] K. Provan, A. Fish, and J. Sydow, "Interorganizational networks at the network level: A review of the empirical literature on whole networks," Journal of Management, vol. 33, no. 3, 2007, pp. 479.

[3] H. Milward, and J. Raab, "Dark networks as organizational problems: Elements of a theory," International Public Management Journal, vol. 9, no. 3, 2006, pp. 333-360.

[4] J. Ayling, "Criminal organizations and resilience," International Journal of Law, Crime and Justice, vol. 37, no. 4, 2009, pp. 182.

[5] D. Lazer, A. Pentland, L. Adamic, S. Aral, A. Barabasi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, and M. Gutmann, "Life in the network: the coming age of computational social science," Science (New York, NY), vol. 323, no. 5915, 2009, pp. 721.

[6] Y. Huang, M. Zhu, J. Wang, N. Pathak, C. Shen, B. Keegan, D. Williams, and N. Contractor, "The Formation of Task-Oriented Groups: Exploring

Combat Activities in Online Games," Proceedings of IEEE SocialCom 2009, pp. 122-127.
[7] D. Huffaker, W. Jing, J. Treem, M.A. Ahmad, L. Fullerton, D. Williams, M.S. Poole, and N. Contractor, "The Social Behaviors of Experts in Massive Multiplayer Online Role-Playing Games," Proceedings of IEEE SocialCom 2009, pp. 326-331.
[8] D. Williams, "The mapping principle and a research framework for virtual worlds," Communication Theory, 2010.
[9] E. Castronova, D. Williams, C. Shen, R. Ratan, L. Xiong, Y. Huang, and B. Keegan, "As real as real? Macroeconomic behavior in a large-scale virtual world," New Media & Society, vol. 11, no. 5, 2009, pp. 685.
[10] E. Castronova, *Synthetic worlds: The business and pleasure of gaming*, University of Chicago Press, 2005.
[11] R. Heeks, "Current Analysis and Future Research Agenda on "Gold Farming: Real World Production in Developing Countries for the Virtual Economies of Online Games," Institute for Development Policy and Management, University of Manchester, 2008.
[12] J. Dibbell, *Play Money: Or, How I Quit My Day Job and Made Millions Trading Virtual Loot*, Basic Books, 2006.
[13] E. Castronova, Synthetic Worlds: The Business and Culture of Online Games, University of Chicago Press, 2005.
[14] G. Lastowka, "ID theft, RMT & Lineage," TerraNova, 2006; http://terranova.blogs.com/terra_nova/2006/07/id_theft_rmt_nc.html.
[15] E. Castronova, "A cost-benefit analysis of real-money trade in the products of synthetic economies," Info, vol. 8, no. 6, 2006, pp. 51-68.
[16] T. Lehtiniemi, "How big is the RMT market anyway?," Virtual Economy Research Network, no. March 2, 2007.
[17] J. Dibbell, "Owned!: Intellectual property in the age of dupers, gold farmers, eBayers, and other enemies of the virtual state," The State of Play: Law, Games, and Virtual Worlds, J.M. Balkin, and B.S. Noveck eds., New York University Press, 2003.
[18] F.G. Lastowka, and D. Hunter, "Virtual Crime," The State of Play: Law, Games, and Virtual Worlds, J.M. Balkin, and B.S. Noveck eds., New York University Press, 2006.
[19] M. Consalvo, Cheating: Gaining advantage in videogames, The MIT Press, 2007.
[20] Tyren, "World of Warcraft Accounts Closed Worldwide," 2006; http://forums.worldofwarcraft.com/thread.html?topicId=59377507.
[21] N. Yee, "Buying gold," Daedalus Project, 2005; http://www.nickyee.com/daedalus/archives/pdf/3-5.pdf.
[22] C. Morselli, C. Giguère, and K. Petit, "The efficiency/security trade-off in criminal networks," Social Networks, vol. 29, no. 1, 2007, pp. 143-153.
[23] B. Erickson, "Secret Societies and Social Structure," Social Forces, vol. 60, 1981, pp. 188.
[24] W. Baker, and R. Faulkner, "The social organization of conspiracy: Illegal networks in the heavy electrical equipment industry," American Sociological Review, vol. 58, no. 6, 1993, pp. 837-860.
[25] C. Morselli, and K. Petit, "Law-Enforcement Disruption of a Drug Importation Network," Global Crime, vol. 8, no. 2, 2007, pp. 109 - 130.
[26] N. Dorn, L. Oette, and S. White, "Drugs importation and the bifurcation of risk: capitalization, cut outs and organized crime," British Journal of Criminology, vol. 38, no. 4, 1998, pp. 537.
[27] M.A. Ahmad, B. Keegan, J. Srivastava, D. Williams, and N. Contractor, "Mining for Gold Farmers: Automatic Detection of Deviant Players in MMOGs," Proceedings of IEEE Social Computing 2009, pp. 340-345.
[28] C. Morselli, *Inside Criminal Networks*, Springer, 2008.
[29] S. Wasserman, and K. Faust, "Centrality & Prestige," Social Network Analysis: Methods and Applications, Cambridge University Press, 1994, pp. 169-219.
[30] A.-L. Barabási, and R. Albert, "Emergence of Scaling in Random Networks," Science, vol. 286, no. 5439, 1999, pp. 509-512.
[31] R. Albert, and A.-L. Barabási, "Statistical mechanics of complex networks," Reviews of Modern Physics, vol. 74, no. 1, 2002, pp. 47.
[32] L. Amaral, A. Scala, M. Barthelemy, and H. Stanley, "Classes of small-world networks," Proceedings of the National Academy of Sciences of the United States of America, vol. 97, no. 21, 2000, pp. 11149.
[33] M. Newman, "Mixing patterns in networks," Physical Review E, vol. 67, no. 2, 2003, pp. 26126.
[34] M. Newman, "Assortative mixing in networks," Physical Review Letters, vol. 89, no. 20, 2002, pp. 208701.
[35] M. Newman, and J. Park, "Why social networks are different from other types of networks," Physical Review E, vol. 68, no. 3, 2003, pp. 36122.
[36] R. Albert, H. Jeong, and A. Barabási, "Error and attack tolerance of complex networks," Nature, vol. 406, no. 6794, 2000, pp. 378-382.
[37] M. Newman, "The structure and function of complex networks," SIAM review, vol. 45, no. 2, 2003, pp. 167-256.
[38] J. Qin, J. Xu, D. Hu, M. Sageman, and H. Chen, "Analyzing terrorist networks: A case study of the global salafi jihad network," Lecture Notes in Computer Science, vol. 3495, 2005, pp. 287-304.
[39] N. Coles, "It's Not What You Know–It's Who You Know That Counts. Analysing Serious Crime Groups as Social Networks," British Journal of Criminology, vol. 41, no. 4, 2001, pp. 580.
[40] G. Robins, "Understanding individual behaviors within covert networks: the interplay of individual qualities, psychological predispositions, and network effects," Trends in Organized Crime, vol. 12, no. 2, 2009, pp. 166-187.
[41] K. von Lampe, and P. Johansen, "Criminal Networks and Trust: On the importance of expectations of loyal behaviour in criminal relations," Organised Crime, Trafficking, Drugs, pp. 102.
[42] B. Coriat, and G. Dosi, "Learning how to govern and learning how to solve problems: On the co-evolution of competences, conflicts and organizational routines," The dynamic firm: the role of technology, strategy, organization and regions, 1998, pp. 103–133.
[43] M. Sparrow, "The application of network analysis to criminal intelligence: An assessment of the prospects," Social Networks, vol. 13, no. 3, 1991, pp. 251-274.
[44] K. Faust, and J. Skvoretz, "Comparing networks across space and time, size and species," Sociological Methodology, vol. 32, 2002, pp. 267-299.
[45] N. Contractor, S. Wasserman, and K. Faust, "Testing Multitheoretical, Multilevel Hypotheses about Organizational Networks," Academy of management review, vol. 31, no. 3, 2006, pp. 681.