Child sexual abuse material networks on the darkweb: a multi-method approach

Bruggen, M. van der

CHAPTER 6

CHARACTERIZING KEYPLAYERS IN CSAM NETWORKS ON THE DARKWEB

This chapter has been published as:

Fonhof, A., Van der Bruggen, M., & Takes, F. (2018). Characterizing keyplayers in child exploitation networks on the dark net. *Complex Networks, 813*, 412-423. https://doi.org/10.1007/978-3-030-05414-4_33

Abstract
This paper studies online child exploitation networks in which users communicate about illegal child pornography material. Law enforcement agencies are extremely interested in better understanding these networks and their keyplayers. We utilize unique real-world network datasets collected from two different online discussion forums on the dark net. Our study of the network structure underlying these forums results in three contributions. First, we propose an approach to identify keyplayers using various centrality measures, allowing us to automatically rank users. Experiments show that our method closely resembles a network-agnostic ranking of users created by domain experts. Second, network metrics are able to characterize a large portion of the users, allowing us to distinguish between regular users, managers and technical moderators. Finally, analyzing the structural properties and distributions of these networks in both the one-mode and two-mode perspective reveals various interesting network-driven insights, such as anti-lurker and anti-law enforcement policies and new user application guidelines. In addition, we found that active users form an elite that participate in more specialized discussions.

6.1 Introduction

Child pornography can be defined as any visual depiction of sexually explicit conduct involving a minor. It has serious damaging effects on the victims and can be considered one of the key social security problems, especially in today's digital society. With the emergence of the dark net (a part of the internet that requires specific software or authorization, see: Egan, 2018) the access to child pornography has been made more secure and anonymous, resulting in growing numbers of users. These factors have made it increasingly valuable to have the right methods and techniques that help agencies to prioritize their law enforcement activities.

In this paper we study data originating from online discussion forums on the dark net. Offenders use these forums to distribute and communicate about illegal child abuse material (Van der Bruggen & Blokland, 2018). Such platforms are often moderated and organized in a professional manner, serving hundreds of thousands of users. To efficiently coordinate law enforcement activities, it is important to target keyplayers that are vital to the existence of these forums, such as administrators, technical moderators and abusers. In this paper we will explore this data as a network and attempt to automatically identify these keyplayers and their role using various network science methods and techniques (Barabási, 2016).

A forum consists of topics and allows multiple users to respond on a certain topic by placing a message, also called a post. This activity can be modelled using a two-mode *topic-to-user* network. To also observe the direct social relationships emerging on these forums, we also project our network to a *user-to-user* network. This allows a number of methods to better understand the network structure to be applied. However, a lot of information in the two-mode network is lost after projection, e.g. how many topics two users responded to or how big a certain topic is that connects two users. Projection can also lead to one-mode network properties that are a result of the projection process rather than the underlying social structure, for example a topic that is linked to many users, which we call a 'big linker'. In the considered child exploitation networks, if two users are linked because they commented on a big generic 'Introduce yourself'-topic, this link is of less significance than if they commented on a specific abuse-related topic. Therefore, we explore different types of projection algorithms in an attempt to select the method which best recreates the forum's underlying social structure.

This paper provides three contributions. First, we study the social structure of the child exploitation networks in an attempt to understand their functioning and governance structure. Second, we analyze the effect of various projection methods on the structure of this network, enabling the third contribution: automated identification of keyplayers and the characterization of user roles.

The rest of this paper is structured as follows. In Section 6.2 we formulate the notions and algorithms that are used in this paper. Related research is discussed in Section 6.3. We describe the data in Section 6.4. Then the proposed approaches are outlined in Section 6.5. Next, experiments are performed in Section 6.6. Finally, conclusions are drawn and future work is suggested in Section 6.7.

6.2 Preliminaries

In this section we discuss the terminology and network metrics used in this paper, adapting the notation of two-mode and one-mode networks by Latapy et al. (2008). Note that there is some minor overlap in notation and symbols for two-mode and one-mode networks. In the remainder of the paper, whenever the context requires so, we will explicitly state in which type of network we are considering the metric.

6.2.1 Two-mode networks

A two-mode (or bipartite) network is made up of two different sets of vertices in which ties exist only between vertices belonging to different sets. The two sets of vertices are users and topics and a tie is a comment that a user posts on a topic. A distinction is often made between the two vertex sets based on which set is considered more responsible for tie creation (called the primary or top vertex set) than the other (secondary or bottom vertex set). We denote a two-mode graph as $G = (\top, \perp, E)$ where $\top$ is the set of top vertices, $\perp$ is the set of bottom vertices and $E \subseteq \top \times \perp$ is the set of undirected edges. In the considered child exploitation networks the topics are the top vertex set, the users are the bottom vertex set, and links denote a user commenting on a certain topic. We will denote the number of top and bottom nodes as $n_\top = |\top|$ and $n_\perp = |_\perp|$, respectively and the total number of nodes is denoted $n = n_\top + n_\perp$. We denote the number of edges as $m = |E|$. We can then define the top and bottom average degree as $k_\top = (m/n_\top)$ and $k_\perp = (m/n_\perp)$, respectively. The total average degree is defined as $\langle k \rangle = (2m/(n_\top + n_\perp)$.

6.2.2 One-mode networks

After projection (which we discuss in Section 6.5.1) we have a one-mode social network $G = (V, E)$ in which users (so, $V = \perp$) are connected by a set of social interaction edges E. An edge denotes that two users replied to the same topic, possibly with an edge weight, dependent on the employed projection method (see Section 6.5.1). We again denote the number of nodes as $n = |V|$ and the edge count as $m = |E|$, and use $k(v)$ for the degree of a node v. The average degree $\langle k \rangle = \frac{1}{n}\sum_{v \in V} k(v)$ is computed by averaging the degree over all nodes. The distance $d(u, v)$ is the length of a shortest path between two nodes u and v. The average path length over all node pairs is denoted by $\langle d \rangle$.

The diameter d_{max} is the maximum distance over all node pairs. The degree assortativity coefficient r describes how nodes are preferentially connected based on their degree, and is defined as:

$$r = \frac{\sum_i e_{ii} - \sum_i a_i b_j}{1 - \sum_i a_i b_j}$$

Here, $a_i = \sum_j e_{ij}$ and $b_j = \sum_i e_{ij}$, where e_{ij} is the fraction of edges from a node with degree i to a vertex with degree j. When $r > 0$ the network is said to be assortative and when $r < 0$ it is disassortative. The average neighbor degree is the average degree of a node's neighbors. The weighted degree of a node v is the sum of edge weights of nodes adjacent to v. The average weighted degree connectivity is the average nearest neighbor weighted degree of nodes with weighted degree k. For further details of these metrics, the reader is referred to Barabási (2016).

6.3 Related work

Below we briefly discuss related work on criminal networks, child exploitation networks and methods to identify keyplayers.

In a recent study, Van der Bruggen and Blokland (2021) discuss child pornography on the internet, arguing that child exploitation networks on the dark net could be classified as criminal organizations. In Duijn (2016), the use of network analysis for detecting and disrupting criminal networks was investigated, yet, child exploitation networks remained beyond the scope. Westlake et al. (2011) studied child exploitation networks, in particular as a connection between websites. They identified the most important website to target for law enforcement based on a metric called network capital.

Latapy et al. (2008) introduce a set of metrics to capture properties of interest in two-mode networks. They provide an alternative to the projection approach, emphasizing that (weighted) projection approaches also produce compelling insight and that the two approaches should be used in an interdependent manner to thoroughly understand the properties of two-mode networks. In this paper, we will evaluate these claims on our two-mode child exploitation network data.

Identifying keyplayers is typically done using centrality measures. More detailed methods for finding keyplayers in social networks, such as those suggested by Borgatti (2006) are aimed at specific tasks, such as optimally transmitting a message through the network or fragmentation of the network by the removal of certain users. However in this paper, we aim to rank the entire set of network users.

6.4 Data

We use two different network datasets, referred to anonymously as dataset A and dataset B due to it being law enforcement sensitive data. The data originates from two distinct child exploitation forums on the dark net that have been taken down by law inforcement and are no longer in operation. Below, we discuss the two forums as well as the metadata added by domain experts.

6.4.1 Forum data

Dataset A consists of 14,659 users and spans a 4 year time period from 2010 to 2014. In order to get access to this forum, users had to provide abusive content that had to be verified by admins. It featured a tiered system, meaning that users were given access to special topics if they presented more unique or self-produced material. This allowed users to gain prestige in the network by actively contributing.

Dataset B has 21,257 users, spans 2 years and was in operation from 2015 to 2017. This forum had a standard approach to user registration, giving users access to almost all topics, apart from a few protected topics for producers and administrators.

The remainder of this paper explores these two-mode networks and their properties. For both datasets, users that did not comment on any topic were excluded, which means that the actual number of members may have been much higher than the number of users included in our study. Furthermore, only unique identifiers for users and topics are studied; no textual or image data was included in this study. Further statistics are provided in Section 6.6.1.

6.4.2 Domain-specific node metadata

Researchers at the Dutch National Police performed the so-called Program Identifying Main Targets (PIM) analysis, partially based on Nolker and Zhou (2005). It uses forum conversation analysis, language analysis and TF/IDF metrics of conversation importance to determine membership roles as well as a ranking of keyplayers.

6.4.2.1 PIM membership roles.

Users were divided into four groups (each with their own percentage of occurrence in dataset A and B):

- **Managers** (0.49% of A, 3.52% of B) are responsible for organizing the forum, recruiting and welcoming new members and enforcing rules.
- **Abusers** (0.59% of A, 4.1% of B) communicate extensively about child abuse and share experiences and fantasies with the community, encourage others to commit criminal activities and may also produce material themselves.
- **Technical users** (0.4% of A, 3.14% of B) focus on developing and sharing anonymity software and providing technical support to other users.
- **Embedded users** (98.3% of A, 89.24% of B) are users that do not fall into any of the first three groups.

6.4.2.2 PIM ranking

A numeric metric was devised by domain experts to assign a value to each user determining its importance. It combines the aforementioned forum conversation analysis and TF/IDF metrics of conversation importance. In addition, users who used particular words identified by domain experts to be characteristic for important users, received a higher metric value. In the PIM ranking, users are ordered by this particular metric value. For details, see Nolker & Zhou (2005).

6.5 Approach

First, Section 6.5.1 discusses the different types of projection methods. Then, Section 6.5.2 explains how we propose to identify and characterize key users.

6.5.1 Determining the right projection method

Here, we build on the definitions in Section 6.2, looking in detail at how the projection from a two-mode to a one-mode network assigns weights to the nodes. The $\top$-projection of graph G is denoted $G_\top = (\top, E_\top)$. Two nodes of $\top$ are linked if they share at least one neighbour in the two-mode network, so $E_\top = \{(u, v), \exists x \in \top: (u, x) \in E$ and $(v, x) \in E\}$. The $\perp$-projection is defined analogously. To uncover the underlying social structure of the forums, in this paper we look at the *user-to-user* network and are therefore only interested in the $\perp$-projection. We will study three projection methods:

- **Unweighted projection** creates an unweighted undirected network of users that commented on the same topic at least once. As mentioned in Section 6.1, here the number of topics that two users commented on is lost.
- **Weighted projection** assigns a weight $w_{u,v} = |N(u) \cap N(v)|$ to each edge (u, v) in the projected network denoting the number of common topics u and v commented on, where $N(v)$ is the set of neighbors of v in the two-mode network. This retains information on the number of topics.
- **Newman's collaboration model** (Newman, 2001) assigns a weight as follows:

$$w_{u,v} = \sum_x \frac{\delta_u^x \delta_u^x}{k(x) - 1}$$

  Here, δ_u^x is 1 if $(u,x) \in E$ in the two-mode network, and 0 otherwise. Also, u and v belong to the $\perp$ node set and x belongs to the $\top$ node set. The value of $k(x)$ is the degree of x in the two-mode network and δ_u^x is 1 if node u is linked to node x in the two-mode network, and 0 otherwise.

For the remainder of the paper, for computing distances, we inverse the weights (so, $w_{u,v} = \frac{1}{w_{u,v}}$), so a lower weight indicates more relatedness of users.

6.5.2 Key user characterization

Users that play a significant role in a network supply human capital, which consists of services that are of importance to the survival of the forum network (Duijn, 2016). For example moderating topics, recruiting new users, distributing resources or helping users with technical questions. To identify these users we will focus on largest connected component (see Section 6.2.2) of the one-mode *user-to-user* network, considering the following metrics:

- **Degree centrality** $C_D(v)$: the fraction of nodes that user v is connected to, defined as $C_D(v) = \frac{1}{n-1} k(v)$.
- **Closeness centrality** $C_C(v)$: this metric computes the distance of node v to each other node in the network. To deal with multiple connected components, we use the highly similar harmonic centrality, formally defined as $C_C(v) = \frac{1}{\sum_u d(u,v)}$.
- **Eigenvector centrality** $C_{EV}(v)$: determines the centrality of a node based on how central, or well-connected, its neighbors are, see Bonacich (1987) for details.
- **PageRank** $C_{PR}(v)$: is a metric that ranks nodes based on the likelihood that a random surfer in the network will arrive at that node, see Brin and Page (1998) for details.
- **Betweenness centrality** $C_B(v)$: this measure computes the number of shortest paths that run through a node v. With σ_{st} the number of shortest paths from s to t and $\sigma_{st}(v)$ the number of shortest paths from s to t that pass through vertex v it is defined as $C_B(v) = \sum_{\substack{s \neq v \neq t \in V \\ s \neq t}} \frac{\sigma_{st}(v)}{\sigma_{st}}$.

To rank users, we propose to sort users by either one of the centrality metrics. To identify the role of a user, we compare the centrality measures between the different roles (as defined in Section 6.4.2). Through this approach we furthermore gain insight in the applicability of these metrics in understanding user roles automatically.

6.6 Results

We explore the characteristics of the two-mode networks in Section 6.6.1. Then we compare the different projection methods in Section 6.6.2. Section 6.6.3 builds upon the projected network, investigating our methods of ranking keyplayers. Finally, Section 6.6.4 analyzes the centrality measurements of different user roles.

6.6.1 Network characteristics

Table 6.1 shows basic statistics about the topology of the two-mode networks. Interesting to note is the relatively large number of top nodes ($n_\top$) and the high average degree for the bottom nodes ($k_\perp$) in dataset A.

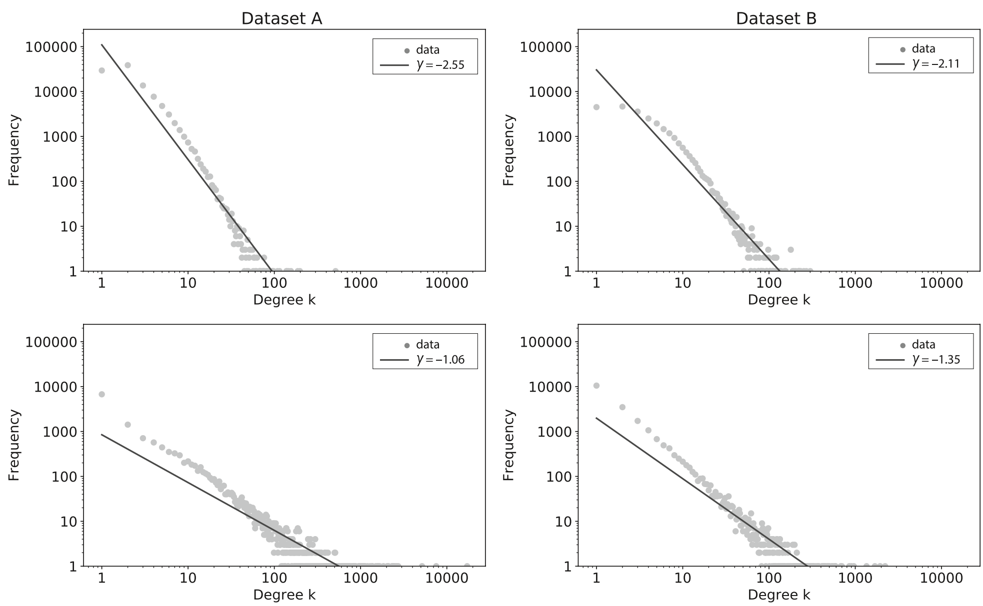
Table 6.1 Two-mode network statistics of the two datasets

	Dataset A	Dataset B
Nodes n	119,742	46,313
Topics $n_\top$	105,083	25,056
Users $n_\perp$	14,659	21,257
Posts m	309,716	145,086
Average degree $\langle k \rangle$	5.2	6.3
Average topic degree $k_\top$	3.0	5.8
Average user degree $k_\perp$	21.1	6.8

6.6.1.1 Degree distributions

The above mentioned difference in average degree can be understood by looking at the degree distribution, shown in Figure 6.1. In real-world data, these distributions follow a power law with exponent γ (Barabási, 2016). We observe that that the top nodes (topics) of dataset A have a higher value of γ than in dataset B, which may indicate that there is more emphasis on smaller topics in dataset A. For the bottom nodes (users) we see the opposite pattern; in dataset A, γ is lower than in dataset B. This indicates that a few users in dataset A comment on almost all topics, which explains the relatively high average value of $k_\perp = 21.1$ in dataset A.

Figure 6.1 Degree distribution of topics (first row) and users (bottom row)

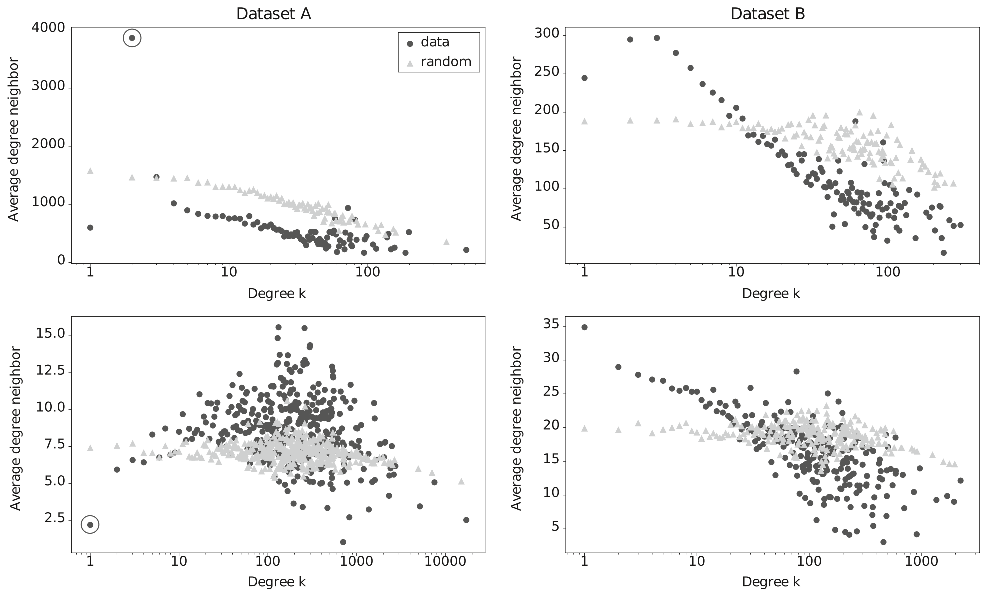


6.6.1.2 Average degree connectivity

Figure 6.2 shows the average degree connectivity (see Section 6.2) for both topics and users. For reference we also plot this value for a random two-mode network with the same degree sequence, generated using the configuration model. The figure highlights various interesting phenomena.

First, the top nodes in dataset B display a negative trend, suggesting that larger topics are commented on by users that are on average less active. However, as shown in Figure 6.2, the bottom nodes (users) do not show this negative correlation. In fact, the most active users comment on topics with relatively few comments. Second, in dataset A topics with 2 comments are on average commented on by users with an average degree of almost 4,000 (highlighted top left node). According to the degree distribution shown in Figure 6.1 there are only a handful of users with a degree larger than 4,000. It turns out that this is the result of the application process of the forum represented by dataset A, where new users had to make an application to gain access, which had to be approved by an administrator (with the high average degree). The bottom left plot of Figure 6.2, showing the average degree connectivity of users, highlights a similar phenomenon occurring at degree $k = 1$ with an average degree connectivity of 2. This represents users who commented once and did so on a topic with only 2 comments; again the aforementioned application topic. This can essentially be seen as an anti-lurker or anti-law enforcement policy; access to the forum is restricted to those who do not provide content. The bottom row of the figure shows that in dataset A there are relatively high post counts for high degree users, indicating that users keep participating and contributing content over time. Altogether, this demonstrates the existence of the tiered system in dataset A, mentioned in Section 6.4, where users providing content were given access to more specialized topics. This pattern was not found in dataset B, which indeed also did not have this policy.

Figure 6.2 Average degree connectivity of topics (first row) and users (bottom row)



6.6.2 Comparing projections

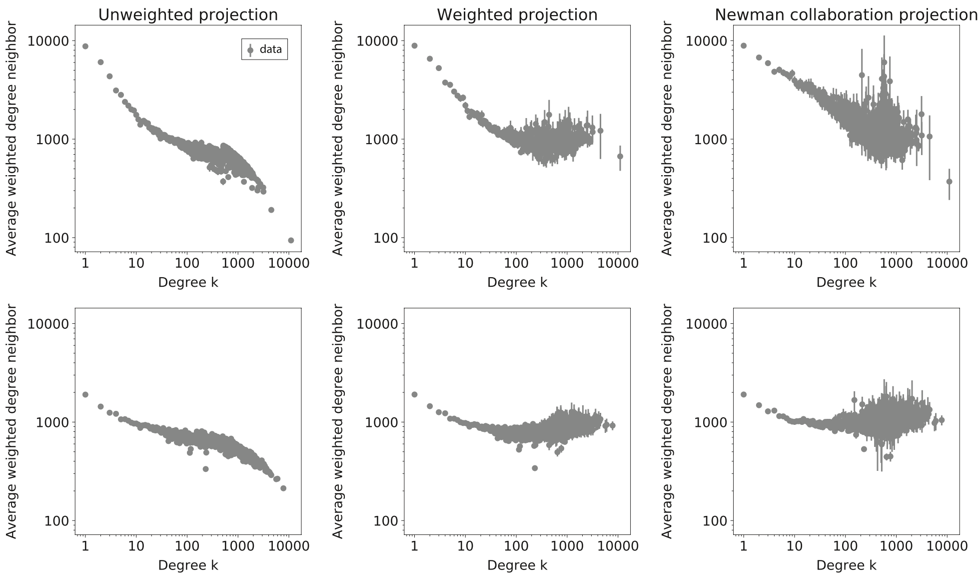
To assess which projection method works best, Table 6.2 shows the degree assortativity coefficient *r* for each of the discussed projection methods. It is highest in both datasets for the weighted projection and Newman collaboration projection. In these two projections, the weighted degree of a node is now scaled by the strength of its relationships. This could imply that using weights better encompasses the phenomenon that users connect with users who are also well connected.

Table 6.2 Degree assortativity coefficient of the projected graphs

	Dataset A			Dataset B		
	Unweighted	Weighted	Newman	Unweighted	Weighted	Newman
r	–0.137	–0.123	–0.038	–0.109	–0.076	–0.049

We further examine this by plotting the average weighted degree connectivity (see Figure 6.3). The leftmost plot of unweighted projection, shows for both datasets a weak negative trend, which may imply that the more users you are connected with the less well connected your neighbors are. However, when we add weights to the edges this effect disappears. This explains the increasing *r* in Table 6.2; it puts additional emphasis on a user that is talking to other important users. Through discussions with domain experts we validated that well connected users are indeed more inclined to chat with other well connected users. In other words: the more important a member is for the community, the more well-connected he is and the more likely to communicate with other well connected users. This can be seen as the existence of an elite of active users. All in all, these experiments suggest that adding weights (either with weighted projection or Newman projection) gives a more accurate representation of the underlying social structure of users.

Figure 6.3 Average weighted degree connectivity of the three projected graphs. The first row shows dataset A and the second row shows dataset B



6.6.3 Identifying keyplayers

The ranking of users based on centrality metrics is influenced by the type of projection used. Therefore, for each of the five centrality metrics in Section 6.5.2 we compute for each of the three projection methods in Section 6.5.1 how well it reproduces the aforementioned PIM ranking (see Section 6.4.2). To do so, we compute the Spearman rank-order correlation coefficient (Ziegel, 2001) between the PIM ranking and the considered centrality measure. A value close to 1 (or -1 for a negative correlation) indicates more agreement on the ordering, whereas a value of 0 means that the ranking are unrelated.

Table 6.3 shows the rank-order correlation coefficients. For both datasets, the weighted projection method outperforms the other two types of projection in recreating the PIM ranking. This is a second piece of evidence suggesting that weights need to be incorporated in the projection step in order to capture important patterns in user interaction. In dataset A we see as much as a rank-order correlation of 0.79 when we use closeness centrality with a weighted projection. In dataset B we have a rank correlation 0.57 with PageRank and the weighted projection method. This shows how a global ranking based on centrality is able to accurately reproduce the PIM ranking generated by domain experts.

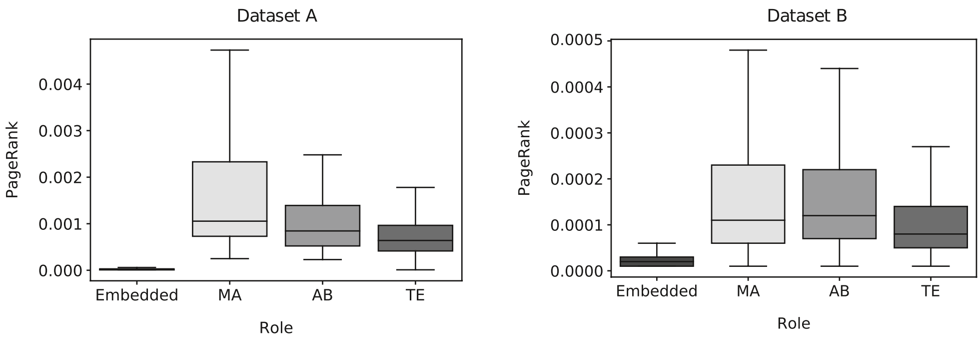
Table 6.3 Spearman rank correlation between PIM ranking and centrality metrics

	Unweighted					**Weighted**					**Newman**				
	C_B	C_C	C_{EV}	C_D	C_{PR}	C_B	C_C	C_{EV}	C_D	C_{PR}	C_B	C_C	EV	C_D	PR
Dataset A	0.5	0.46	0.35	0.44	0.47	0.57	0.79	0.75	0.44	0.76	0.52	0.77	0.76	0.44	0.75
Dataset B	0.54	0.43	0.43	0.51	0.53	0.55	0.53	0.46	0.51	0.57	0.53	0.44	0.42	0.51	0.55

6.6.4 User roles

The goal of user role identification is to find network-driven characteristics that help understand the position of particular groups of users. In the subsection above we found that PageRank had the highest rank-order correlation in dataset B, and the second highest in dataset A. Figure 6.4 shows the differences between the user roles (see Section 6.4.2), again based on PageRank. We observe that users with a special role (managers (MA), abusers (AB) and technical users (TE)) consistently have a more central position than embedded (regular) users. Furthermore, managers and abusers are in turn more central than technical users, which can be explained by the fact that these keyplayers are more individualistic and focus more on building applications, mainly talking with their peers. Managers on the other hand have to communicate with all users and are thus more central.

Figure 6.4 Distribution of PageRank values with weighted projection for different user roles (see Section 6.4.2)



6.7 Conclusion and future work

In this paper we studied two unique network datasets from online child exploitation forums, in which users comment on certain child abuse related topics. Through analyzing the degree distribution of the two-mode network we found that larger topics are commented on by relatively less active users, likely because these topics are more easily found and talk about a more easily accessible subject. Topics with few comments were commented on by relatively more active users, hinting at the existence of an elite of users contributing to more specialized discussions, containing keyplayers with roles important to the forum. The average degree connectivity of dataset A revealed the admission procedure of this forum, where users had to provide content in order to gain access to the forum. We also discovered that using a weighted form of projection to obtain the one-mode network is crucial for obtaining the network's underlying social structure of the user-to-user network. Using weighted projection and closeness centrality we were able to obtain a high (up to 79% for dataset A) rank-order correlation coefficient with a ranking of users generated by domain experts. This demonstrates the power of network metrics in identifying keyplayers. Finally, we evaluated the different user roles, and found that it is possible to distinguish between regular users and keyplayers based on their centrality values. It furthermore revealed the more individualistic role of technical users dealing with the forum setup, encryption and maintenance. In general, the proposed network approach has a number of advantages over content-based analysis of dark net forum data, as it only requires the structure of the forum, and as such can deal with encrypted posts and forums in unknown foreign languages.

In future work, we want to investigate if we can devise a classification model to accurately determine the role of a given user, building upon the results of the analysis and characterization of user roles through centrality. Furthermore, we want to investigate private messages sent on the forums in an attempt to understand the extent to which the observed social interaction in the projected networks captures direct social interaction in the network.