# Predicting Behavioural Patterns in Discussion Forums using Deep Learning on Hypergraphs

Arya, D.; Rudinac, S.; Worring, M.

**DOI**
10.1109/CBMI.2019.8877384

**Publication date**
2019

**Document Version**
Final published version

**Published in**
2019 International Conference on Content-Based Multimedia Indexing (CBMI)

# Predicting Behavioural Patterns in Discussion Forums using Deep Learning on Hypergraphs


Devanshu Arya
*University of Amsterdam*
Amsterdam, The Netherlands
d.arya@uva.nl

Stevan Rudinac
*University of Amsterdam*
Amsterdam, The Netherlands
s.rudinac@uva.nl

Marcel Worring
*University of Amsterdam*
Amsterdam, The Netherlands
m.worring@uva.nl



***Abstract*—Online discussion forums provide open workspace allowing users to share information, exchange ideas, address problems, and form groups. These forums feature multimodal posts and analyzing them requires a framework that can integrate heterogeneous information extracted from the posts, i.e. text, visual content and the information about user interactions with the online platform and each other. In this paper, we develop a generic framework that can be trained to identify communication behavior and patterns in relation to an entity of interest, be it user, image or text in internet forums. As the case study we use the analysis of violent online political extremism content, which has been a major challenge for domain experts. We demonstrate the generalizability and flexibility of our framework in predicting relational information between multimodal entities by conducting extensive experimentation around four practical use cases.**



***Index Terms*—hypergraphs, geometric deep learning, discussion forums, semantic concept detection, entity linking**


## I. Introduction

A large amount of visual and textual content is posted daily in different social networking and content sharing platforms, where users can express their thoughts and share experiences. Pervasive nature of internet and social media has not only made it possible to communicate and demonstrate radical views and intentions, but also to connect to other persons with similar interests. Due to this high reachability and popularity of social media, people also use these platforms for planning events and mobilizing others for protests, public demonstrations, promoting violent extremist ideologies, and spreading racist opinions. The problem of automatic identification of such online radicalization and prediction of social unrest is of paramount importance for law enforcement agencies. It requires collection, fusion and analysis of 'weak signals' or 'digital traces' which are present on social media. Current analysis techniques focus mostly on hashing and filtering of known extremist multimedia items. However, aiding domain experts in rigorous large-scale empirical analysis requires novel multimodal tools designed to handle unstructured data from diverse information channels, especially internet discussion forums.

Discussion forums are a type of social multimedia network where people can meet, form groups, discuss common interests and exchange ideas. Through the use of discussion forums, it is also possible for members of the public, whether supporters or detractors of a group, to engage in debate. This may assist the terrorist group in adjusting their position and tactics and, potentially, increasing their levels of support and general appeal [1]–[3]. This is also noted in Europol's annual terrorism situation and trend report for 2012, which warns that internet forums present effective means for addressing target audiences, and "recruiting" supporters with no off-line links to terrorist organizations [4]. By just analyzing the content after it has been shared in "extremosphere" of these forums can lead to a delay in detecting critical events, which can prove to be a massive loss in the future. Hence, an effective framework is needed for predicting future communication behaviour between users or communities within the (often implicit) social networks hosted by these forums. Major challenge in developing such framework is the presence of low quality content in contextual metadata and the large volume of information in internet forums.

In this paper, we construct a pipeline, as shown in figure 1, which can be used to identify communication behavior and patterns in violent online extremism forums. Our framework is built upon the methodology developed in [5] where the authors presented an approach to predicting links and groups between entities (which can be images, users, posts, groups etc.) within a social multimedia network such as Flickr. An entity can be any of the visible constituents of a post in social network. For example, in Instagram, entities consist of tags, image, video, user, location and caption. In this work, we extend the methodology proposed in [5] for the use with heterogeneous entities in discussion forums, which contain more unstructured information. It can enhance the Law Enforcement Agencies (LEAs) to exploit future interactions between entities within a network, for instances: (a) *mob formation* - which type of people or forums a particular user(s) might be interested in to interact with, (b) *deciphering hidden messages from an image* - what kinds of images can be associated with a post or (c) *content classification* - which type of category a post might belong to. These use cases requires analysis of posts at a particularly high semantic level. Hence, we focus on extracting semantic concepts, such as topics, personages, locations and gender from text using entity linking and visual concepts (such as TRECVID [6] and ImageNet [7] concepts) from images and videos. The usage of such semantic concepts is important since


978-1-7281-4673-7/19/$31.00 ©2019 European Union

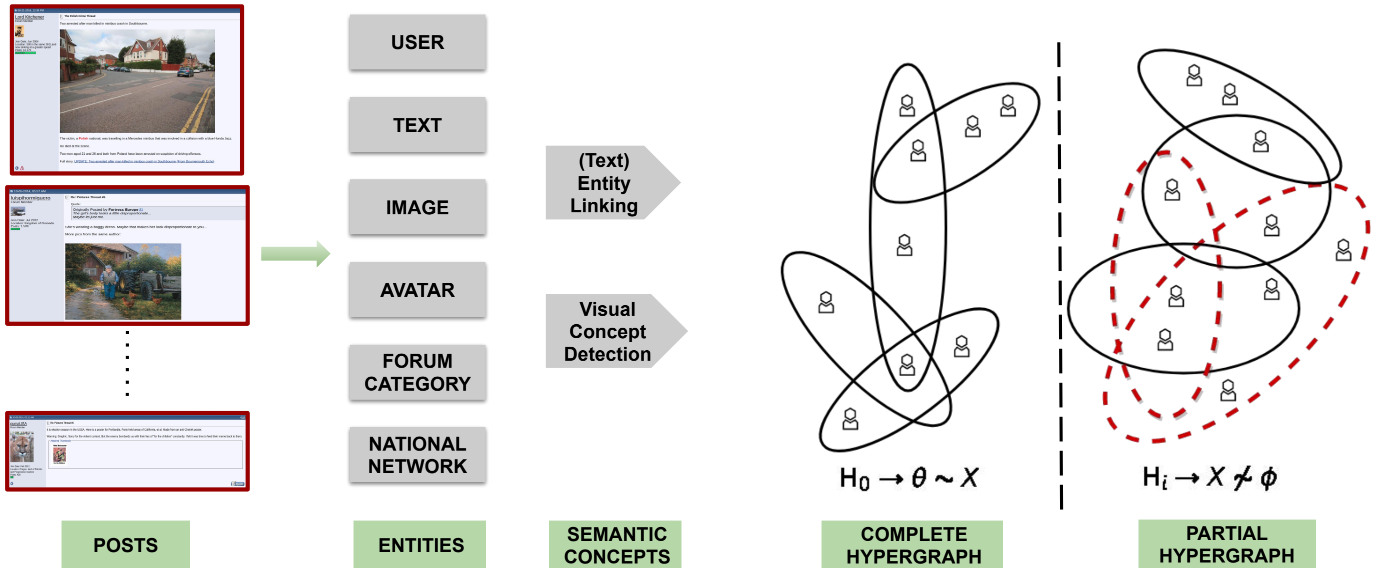



Fig. 1: An example of the proposed pipeline of the framework. It represents a typical post with multimodal entities in Stormfront, a white nationalist, white supremacist and neo-Nazi Internet forum followed by the extraction of semantic concepts from the post and then construction of a hypergraph framework using the entities and concepts as proposed in [5]. $H_0$ and $H_i$ represents complete and partial hypergraph. The goal is to predict missing information i.e. generate hyperedges (red-dotted line) on $H_i$.

the results are intended to be interpretable by the end users.

To facilitate such applications, we extract semantic (visual and text) concepts from the posts and then train a model based on relational information between entities in the social network. These relations can either be formed between entities of different modality or between entity and its metadata information. Uncovering hidden relations between multimedia items has long been a topic of research in multimedia information retrieval. Graph-based approaches have been prominent for their ability to represent and analyze such problems. Recent works [8] [9] [10] on geometric deep learning aim at formulating convolutional neural networks to data represented on graphs. The key idea in geometric deep learning is to devise a method for representation learning that can capture structural information within non-Euclidean domains, especially graphs. However, there are two major challenges for their wider adoption in learning multimodal relations in discussion forums. Firstly, graph-based approaches are hindered by the challenges related with associating and learning semantic concepts due to the presence of unstructured textual data and low quality images. In addition, inefficient representation of multimedia post can lead to loss of available information or capture it only partially.

The main contributions of the paper are:

- We present a framework, which combines semantic concepts and contextual relations between entities in a discussion forum to predict communication behavior and patterns in relation to various types of entities.
- We demonstrate the flexibility of such framework in tackling a variety of potential use cases arising during the analysis of violent political extremism forums. Multimedia analysis is further shown effective in aiding the domain experts involved in the qualitative analysis of these forums.
- Our experiments provide insights into the usefulness of the relations between individual modalities and semantic features, which can be exploited to unravel implicit information about diverse entities in a discussion forum.

The remainder of this paper is organized as follows. In Section 2 we provide an overview of related work. Then in Section 3 we introduce our approach and in sections 4 and 5 we present the experimental setup and results. Section 6 concludes the paper.

## II. Related Work

This section describes and discusses related work on the methods for learning patterns and behaviours of entities in a social network.

Understanding how users behave when they connect to social networking sites creates opportunities for richer studies of social interactions, better detection of irregular behavior and improved design of content distribution systems. Jin et al. [11] presented an elaborate survey on the importance of analysis and characterization of user behaviours in online social networks, highlighting the different perspectives that are shaping the ongoing work in the field. The need for analysis of user behaviors have now become even more interesting with the rise of social multimedia network. The presence of multimodal entities in social network, has shifted the primary focus of multimedia community to go beyond the structural analysis of network [12] [13] and towards the analysis of content at a higher semantic level [14].

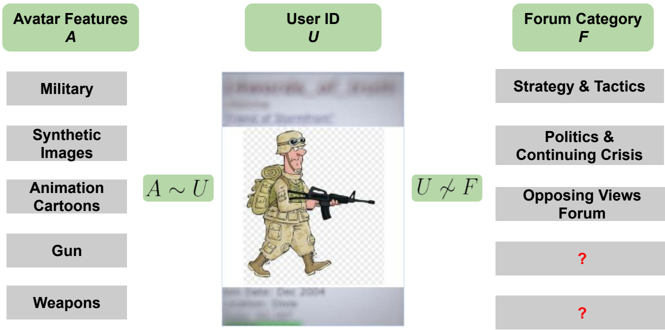


Fig. 2: An example from Use Case 4, where the known information for a user is his/her User ID ($U$), Avatar Features ($A$) and partially known Forum Categories that he/she posted. The goal is to predict these unknown forum categories using the relations $A \sim U$ and $U \not\sim F$. Similar, examples can be drawn from the other use cases.

Moreover, the advantages of combining semantic concepts with the network structure has made graph-based approaches very popular. The application of graph-based methods on social network varies from link prediction [15], discovering social circle in ego networks [16], music recommendation by combining social media information and music content [17] to categorizing violent online political extremism content [18]. Hypergraphs [19], in particular, were proven to be highly efficient in capturing relational information in multimodal social networks [5] [20] [21].

## III. Approach

Given a multimodal post from an arbitrary discussion forum, our goal is to construct a framework that can encode semantic concepts about an entity in the form of relations, and then predict valuable information about them. Our framework can, in general, allow a user to perform either of two tasks: (1) predict implicit relations between multimodal entities or (2) extract additional semantic concept about entities. We combine these two tasks by using a common graph-based approach that can provide users with the flexibility to train a model according to their usage. Below we describe the three main facets of our formulation - representation, information flow and generalizability.

### A. Representation

We represent information in a multimodal post using hypergraphs due to their numerous advantages over traditional graph based representation, one of them being efficient capture of higher-order relations [5] [22]. A hypergraph is a generalization of the graph in which the edges, called hyperedges, are arbitrary non-empty subsets of the vertex set and may therefore connect any number of vertices. The nodes are kept the same as in a graph but a hyperedge can connect even all the nodes at once as compared to a traditional graph where an edge is always a connection between 2 nodes. In particular, a set of multimodal entities in a social network can be viewed as a hypergraph whose vertices are the individuals and whose hyperedges are the common properties between them [20].

The other advantage of using hypergraphs is the ease of modifying definitions of nodes and hyperedges. In this work, we will exploit this property in order to merge the two tasks mentioned in section III. For both tasks, we construct a hypergraph in which its nodes represent the main entity (for which relations/information needs to be predicted) connected through hyperedges which can represent either entities from other modality or metadata information about the main entity. So, the problem reduces to that of generating hyperedges across the main entity which in-turn can represent (a) relations between multimodal entities or (b) relations between metadata and the main entity. In this way, we can devise a learning algorithm which can merge both tasks and is devoid of any loss in available information.

### B. Information Flow

The next challenge is to construct a pipeline extracting entities and concepts from posts to then learn relational information. We aim at encoding information from the available set of relations for an entity to predict the unknown sets of relations. For extracting semantic information from a post, we employ entity linking for text, where the idea is to link the text to an external knowledge base such as Wikipedia [18], [23] and for visual concepts we extract 346 TRECVID semantic concepts [6]. Further, we use a robust model for extracting features and learn relational information [5] from the posts represented on hypergraphs. This model is based on geometric matrix completion solution initially proposed in [24]. We formulate the relation prediction task as a matrix completion problem, where rows and columns represent two separate entities. This matrix is derived from the incidence matrix of partial hypergraph, where the vertices forms the rows and the edges forms the columns. For example, the images posted in discussion forums will be represented on the rows of a matrix (and vertices of hypergraph) while the columns (corresponding edges in hypergraph) can be the forum categories. Thus, each entry of this matrix will have a binary value representing presence/absence of an image in a particular forum category. The aim is to complete this matrix using auxiliary information about images from the complete hypergraph. Hence, to extract combined relational features we use Multi-Graph Convolution Networks (MGCNN). To explain further, we give a brief background about low rank matrix completion and Multi-Graph Convolution Networks.

Low rank matrix completion involves recovering a matrix $M \in \mathbb{R}^{N_1 \times N_2}$ of rank $R << min(N_1, N_2)$ from a subset of its entries $\Omega$. To concisely put, given partial observation of $M$ over an index set $\Omega \subset (1, 2, .., N_1) \times (1, 2, .., N_2)$ the task is to select the matrix with the lowest rank. Let $X$ denotes the matrix to recover and $M_\Omega$ is the set of the known entries. However, rank minimization is an NP-hard optimization problem and in many real world matrix completion problems, the entities defined on rows and/or columns share many common attributes. These entities can thus be encoded using graphs by exploiting their proximity information. Incorporating proximity information forces the solution for matrix completion task to be smooth on these row and column graphs. For the row graph, entities defined in the rows forms the vertices and each row of the matrix can be thought of as signals defined on its vertices. In order to combine information from both the row and column graphs, [25] used the concept of Graph Fourier Transform on matrices. Taking into account that Fourier Transform operation is separable and symmetric, the two dimensional transforms can be computed as sequential row and column one-dimensional transforms. Hence the corresponding Fourier transform of matrix $X$ is given by $\mathcal{F}(X) = \Phi_r^T X \Phi_c$, where $\Phi_r$ and $\Phi_c$ are the eigenvectors of row and column graphs with $\mathbb{L}_r$ and $\mathbb{L}_c$ as the corresponding Laplacian matrices respectively. Further, Monti et.al. [24] proposed Multi-Graph Convolutional Networks (MGCNN) that aims at extracting spatial features from the matrix. Given a matrix $X \in \mathbb{R}^{N_1 \times N_2}$, MGCNN is given by

$$\widetilde{X} = \sum_{j,j'=0}^{q} \theta_{j,j'} T_j(\mathbb{L}_r) X T_{j'}(\mathbb{L}_c) \tag{1}$$

where, $\Theta = \theta_{j,j'}$ is the $(q+1)\times(q+1)$ represents coefficient of filters and $T_j(.)$ denotes the Chebyshev polynomial of degree $j$. Using this equation as the convolutional layer of MGCNN, it produces $q$ output channels ($N_1 \times N_2 \times q$) for matrix $X \in \mathbb{R}^{N_1 \times N_2}$ having a single input channel .

In this way, we extract features and then combine all the information channels in one framework using both the semantic concepts and their contextual relations with the entities.

### C. Generalizability

The proposed framework should be generalizable to different use cases consisting of any form of relevant information. Let $X$ and $\theta$ be the main entity and the known concept/entity respectively. From all the available information, we know all the relations of entities in $X$ to that with $\theta$, let this relation be represented by ($\theta \sim X$). We aim to predict all the relations of $X$ with another type of concept/entity($\phi$) for whom we know partial relations ($X \not\sim \phi$). We define ($\theta \sim X$) and ($X \not\sim \phi$) on two different hypergraphs $H_0$ and $H_i$ respectively and then use the learning model proposed in [5] to complete the partial information in $H_i$. We will represent the complete framework by ($\theta \sim X \not\sim \phi$). Our generic approach representing relations on hypergraph and the learning model make the framework applicable in a variety of settings and use cases. Especially in case of discussion forums, adding any semantic concepts would not alter the pipeline of the proposed framework.

## IV. Experimental Setup

### A. Dataset

We use Stormfront, a white nationalist, white supremacist and neo-Nazi Internet forum as the testbed for this study. The forum contains 40 high-level categories, indicating topics of discussion, ranging from "Politics and Continuing Crises", "Strategy and Tactics" and "Ideology and Philosophy" to the topics relevant to national chapters, e.g. Stormfront en Francais and Stormfront en Espanoly Portugus. Typically, in Stormfront, a user posts a content (text, images, videos, links etc.) in one of the relevant forum categories. These users often have a small avatar image and/or somewhat larger profile picture. Recently, Rudinac et al. [18] deployed graph convolutional neural networks for classifying 2 million user posts from Stormfront. In this paper, we use the same data setup as mentioned in [18] for all the use cases. We include the following items of a post from the dataset for our experiments:

- Post ID (P): Unique ID given to each post
- User (U): User who posted it
- Avatar Features(A): Features extracted from display picture of user's avatar
- Forum Category (F): Type of category in which a post has been shared
- User Topics (T): Topic of interest for a user
- Semantic Entities (E): Relevant entities present in textual content of a post

### B. Use Cases

To demonstrate the effectiveness of our framework we conduct a set of experiments organized around the following use cases:

**Use Case 1: ($E \sim P \not\sim U$)** In the first case, we aim to predict potential users who would be interested in interacting with a certain post. This can help in understanding and tracking communication patterns of certain users. In this case, users are the partial information for the posts while the semantic entities associated with the post will be used for learning the relations between the posts.

**Use Case 2: ($F \sim U \not\sim T$)** [18] the authors conjecture that the user preferences are a good predictor of the post category. We take motivation from this result for the second use case. Often we require more information about a particular user, but due to very limited user activity, it is not possible to extract such information. We conjecture that the community formed by users posting in different forum categories can have sufficient clues which can be exploited to extract more information (user topics) about an arbitrary user.

**Use Case 3: ($E \sim P \not\sim F$)** It is often necessary to categorize posts based on their national network due to the formation of local mob or even event organizations. The extracted semantic entities might carry sufficient information for such classification, as the national chapters are characterised by a certain number of topics, more frequent use of national (i.e. non-English) language, and a closed group of users discussing the matters of regional relevance.

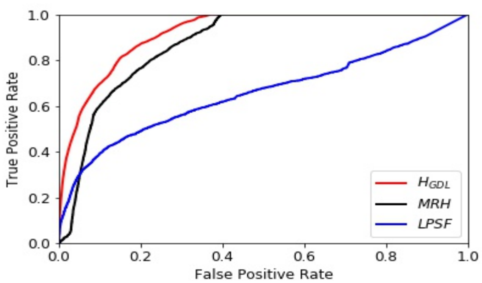
**Use Case 4: ($A \sim U \not\sim F$)** We aim at identifying the properties of avatars specific of a particular post category. This is to investigate which semantic concepts are more commonly appearing in a certain forum categories, for example "For Stormfront Ladies Only" forum as compared to the other forums on Stormfront. The particular use case came from the domain experts investigating the role and portrayal of women

TABLE I: Table showing the data setup of the 4 use cases. For training, the framework uses all the relations $\theta \sim X$ and 40% of $X \sim \phi$. The goal is to predict the rest of $X \sim \phi$ relations.

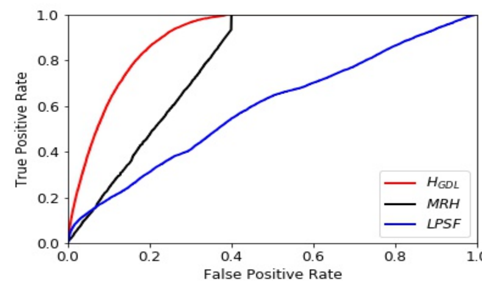
| | **Use Case 1** | **Use Case 2** | **Use Case 3** | **Use Case 4** |
|---|---|---|---|---|
| $\boldsymbol{\theta}$ | $E$: 3,319 | $F$: 39 | $E$: 3,319 | $A$: 290 |
| **X** | $P$: 10,000 | $U$: 10,252 | $P$: 10,000 | $U$: 10, 252 |
| $\boldsymbol{\phi}$ | $U$: 4,218 | $T$: 258 | $F$: 31 | $F$: 39 |
| $\boldsymbol{\theta \sim X}$ | $E \sim P$: 36,640 | $F \sim U$: 40,418 | $E \sim P$: 36,640 | $A \sim U$: 51,260 |
| $\boldsymbol{X \sim \phi}$ | $P \sim U$: 10,000 | $U \sim T$: 29,321 | $P \sim F$: 55,280 | $U \sim F$: 40,418 |

TABLE II: Performance of our approach ($H_{GDL}$ as compared with standard models $MRH$ and $LPSF$

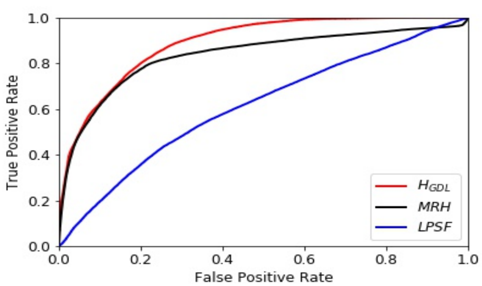
| | **Use Case 1** | | **Use Case 2** | | **Use Case 3** | | **Use Case 4** | |
|---|---|---|---|---|---|---|---|---|
| | AUC | EER' | AUC | EER' | AUC | EER' | AUC | EER' |
| $H_{GDL}$ | 86.4 | 78.5 | 88.7 | 80.3 | 80.6 | 72.5 | 89.2 | 83.0 |
| MRH | 79.1 | 69.9 | 77.8 | 70.2 | 78.4 | 70.6 | 84.6 | 76.5 |
| LPSF | 63.1 | 60.2 | 60.3 | 57.9 | 71.2 | 68.4 | 62.9 | 60.3 |



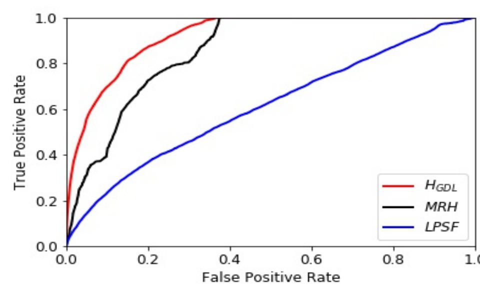

(a) Use Case 1




(b) Use Case 2




(c) Use Case 3




(d) Use Case 4

Fig. 3: Receiver Operating Characteristics (ROC) curve showing the performance of the models on each of the 4 use cases. The hypergraph-based geometric deep learning model ($H_{GDL}$) has significant advantage as compared to other methods on all the 4 use cases.

in right-wing extremist networks. Based on the avatars and profile pictures, we are trying to identify users likely to be associated with a particular, specialised discussion forum (e.g. “dating advice” or “religion”). Figure 2 shows an example of the user’s features extracted from visual concepts and the partial information of his/her forum categories.

### *C. Experiments*

We start our experimental evaluation by showing the performance of our framework on all the 4 use cases. The corresponding number of entities used in each use-cases and the total number of relations formed among them is shown in Table I. As seen from the table, each entity has multiple values in common with other entities, resulting in a multitude of relations. For example, a user ($U$) can post in multiple forum categories about various issues. This will account for a large number of $F \sim U$ and $U \sim T$ relations. For our experimental setup, we randomly sample 40% of these relations and keep them aside to use as a test set. The remaining relations are used to construct the partial hypergraph $H_i$ for training the model.

## V. RESULTS

We report the results of our framework and compare them to hypergraph based algorithm ($MRH$) [17], [20] and a graph-based model trained on social network features ($LPSF$) [12] for the same tasks. $LPSF$ trains a neural network on popular features like Page Rank, Number of Common

Neighbors, Preferential Attachment etc. extracted from a social network. To evaluate the performance of our model and show its advantages over other methods, we plot the Receiver Operating Characteristic (ROC) curves for each task. The ROC curve depicts how well a model is able to predict the presence/absence of any information in an entity. Figure 3 shows the performance of the models on the 4 use cases.

To further quantify the results, we calculate $AUC$(Area Under Curve of the ROC plot) and $EER' = 100\% - EER$ (Equal Error Rate) for all the three methodologies. $EER$ corresponds to the point on the ROC curve that corresponds to an equal probability of miss-classifying a positive or negative sample. Both these numbers are very important indicators of a model's overall performance, the higher the $AUC$ and $EER'$ the higher the accuracy of the system. We show them in Table II, where it can be seen that our model, in general, overperforms alternatives in predicting information of any type of entities.

## VI. Conclusion

In this paper, we constructed a framework which can be used to learn and predict relational information within a discussion forum. The generalizability of the framework provides the flexibility to the end users to formulate their own use cases irrespective of domain-specific constraints. As a test bed for our study we use the analysis of a realistic collection of data from Stormfront, a violent online political extremism forums. The experiments are conducted around research questions raised by the domain experts and demonstrate the effectiveness of our approach in providing implicit information about users of a forum. The results confirm the merit of our approach to geometric deep learning on hypergraphs and suggest that in case of multimodal data, the proposed framework can be used for designing a case study. Finally, on four example use cases we demonstrated that this technique may be a valuable asset to domain experts performing qualitative analysis of violent online political extremism.

## Acknowledgment

This research has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700381 (ASGARD project).

## References

[1] G. Weimann, *www. terror. net: How modern terrorism uses the Internet*. DIANE Publishing, 2004, vol. 116.
[2] R. Gibson and S. Ward, "A proposed methodology for studying the function and effectiveness of party and candidate web sites," *Social science computer review*, vol. 18, no. 3, pp. 301–319, 2000.
[3] M. Conway, "Terrorist's use of the internet and fighting back," *Information and Security*, vol. 19, p. 9, 2006.
[6] C. Snoek, K. Van De Sande, D. Fontijne, A. Habibian, M. Jain, S. Kordumova, Z. Li, M. Mazloom, S. Pintea, R. Tao *et al.*, "Mediamill at trecvid 2013: Searching concepts, objects, instances and events in video," in *NIST TRECVID Workshop*, 2013.
[4] E. TE-SAT, "Terrorism situation and trend report," 2012.
[5] D. Arya and M. Worring, "Exploiting relational information in social networks using geometric deep learning on hypergraphs," in *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval*. ACM, 2018, pp. 117–125.
[7] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. Ieee, 2009, pp. 248–255.
[8] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, "Geometric deep learning: going beyond euclidean data," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 18–42, 2017.
[9] D. K. Duvenaud, D. Maclaurin, J. Iparraguirre, R. Bombarell, T. Hirzel, A. Aspuru-Guzik, and R. P. Adams, "Convolutional networks on graphs for learning molecular fingerprints," in *Advances in neural information processing systems*, 2015, pp. 2224–2232.
[10] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Advances in Neural Information Processing Systems*, 2016, pp. 3844–3852.
[11] L. Jin, Y. Chen, T. Wang, P. Hui, and A. V. Vasilakos, "Understanding user behavior in online social networks: A survey," *IEEE Communications Magazine*, vol. 51, no. 9, pp. 144–150, 2013.
[12] X. Wang and G. Sukthankar, "Link prediction in multi-relational collaboration networks," in *Proceedings of the 2013 IEEE/ACM international conference on advances in social networks analysis and mining*. ACM, 2013, pp. 1445–1447.
[13] Z. Yin, M. Gupta, T. Weninger, and J. Han, "Linkrec: a unified framework for link recommendation with user attributes and graph structure," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 1211–1212.
[14] J. McAuley and J. Leskovec, "Image labeling on a network: using social-network metadata for image classification," *Computer Vision–ECCV 2012*, pp. 828–841, 2012.
[15] C. P. Diehl, G. Namata, and L. Getoor, "Relationship identification for social network discovery," in *AAAI*, vol. 22, no. 1, 2007, pp. 546–552.
[16] J. Leskovec and J. J. Mcauley, "Learning to discover social circles in ego networks," in *Advances in neural information processing systems*, 2012, pp. 539–547.
[17] J. Bu, S. Tan, C. Chen, C. Wang, H. Wu, L. Zhang, and X. He, "Music recommendation by unified hypergraph: combining social media information and music content," in *Proceedings of the 18th ACM international conference on Multimedia*. ACM, 2010, pp. 391–400.
[18] S. Rudinac, I. Gornishka, and M. Worring, "Multimodal classification of violent online political extremism content with graph convolutional networks," in *Proceedings of the on Thematic Workshops of ACM Multimedia 2017*. ACM, 2017, pp. 245–252.
[19] C. Berge, *Graphs and Hypergraphs (North-Holland mathematical library; v. 6)*. Elsevier, 1973.
[20] D. Li, Z. Xu, S. Li, and X. Sun, "Link prediction in social networks based on hypergraph," in *Proceedings of the 22nd International Conference on World Wide Web*. ACM, 2013, pp. 41–42.
[21] Y. Gao, M. Wang, H. Luan, J. Shen, S. Yan, and D. Tao, "Tag-based social image search with visual-text joint hypergraph learning," in *Proceedings of the 19th ACM international conference on Multimedia*. ACM, 2011, pp. 1517–1520.
[22] D. Zhou, J. Huang, and B. Schölkopf, "Learning with hypergraphs: Clustering, classification, and embedding," in *Advances in neural information processing systems*, 2007, pp. 1601–1608.
[23] D. Milne and I. H. Witten, "Learning to link with wikipedia," in *Proceedings of the 17th ACM conference on Information and knowledge management*. ACM, 2008, pp. 509–518.
[24] F. Monti, M. Bronstein, and X. Bresson, "Geometric matrix completion with recurrent multi-graph neural networks," in *Advances in Neural Information Processing Systems*, 2017, pp. 3700–3710.
[25] V. Kalofolias, X. Bresson, M. Bronstein, and P. Vandergheynst, "Matrix completion on graphs," *arXiv preprint arXiv:1408.1717*, 2014.