A

Tracking Illicit Drug Dealing and Abuse on Instagram using Multimodal Analysis

XITONG YANG, University of Rochester
JIEBO LUO, University of Rochester

Illicit drug trade via social media sites, especially photo-oriented Instagram, has become a severe problem in recent years. As a result, tracking drug dealing and abuse on Instagram is of interest to law enforcement agencies and public health agencies. However, traditional approaches are based on manual search and browsing by trained domain experts, which suffer from the problem of poor scalability and reproducibility. In this paper, we propose a novel approach to detecting drug abuse and dealing automatically by utilizing multimodal data on social media. This approach also enables us to identify drug-related posts and analyze the behavior patterns of drug-related user accounts. To better utilize multimodal data on social media, multimodal analysis methods including multi-task learning and decision-level fusion are employed in our framework. We collect three datasets using Instagram and web search engine for training and testing our models. Experiment results on expertly labeled data have demonstrated the effectiveness of our approach, as well as its scalability and reproducibility over labor-intensive conventional approaches.

CCS Concepts: •**Applied computing** → **Sociology;**•**General and reference** → *Empirical studies;*

Additional Key Words and Phrases: multimodal analysis; illicit drug; social media

ACM Reference Format:
Xitong Yang, Jiebo Luo, 2016. Tracking Illicit Drug Dealing and Abuse on Instagram using Multimodal Analysis. *ACM Trans. Intell. Syst. Technol.* V, N, Article A (January YYYY), 15 pages.
DOI: http://dx.doi.org/10.1145/0000000.0000000

1. INTRODUCTION

In recent years, illicit drug use has become a major problem for the society and the trend has been steadily growing. According to the statistics of the annual National Survey on Drug Use and Health (NSDUH) [1], an estimated 24.6 million Americans aged 12 or older (9.4% of the population) had used an illicit drug in the past month in 2013. This number is up from 8.3% in 2002. The increase of illicit drug use reveals the growth of illicit drug trade in the US.

Except for selling their goods offline as the traditional way, drug dealers nowadays use social media sites for drug sales[2]. With a massive number of active users, social media sites have become incredibly effective tools for advertising illegal drugs, especially for the photo-oriented sites like Instagram. For example, thousands of accounts on Instagram are currently selling marijuana, prescription painkillers, and other illicit drugs effectively in an open drug market. One can easily get access to these illegal drugs by a quick hashtag search on Instagram. It has been noticed and reported by law enforcement agencies and public health agencies that illegal drug trade on social me-

[1] https://www.drugabuse.gov/publications/drugfacts/nationwide-trends
[2] http://drugabuse.com/featured/instagram-drug-dealers/

Author's addresses: Xitong Yang and Jiebo Luo, Department of Computer Science, University of Rochester, Rochester, NY 14627.
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© YYYY ACM. 2157-6904/YYYY/01-ARTA $15.00
DOI: http://dx.doi.org/10.1145/0000000.0000000

dia will aggravate drug abuse, especially for teenagers [3] [4]. Therefore, it is imperative to tackle drug dealing and abuse on social media. Given the vast data on social media and limited resources of law enforcement agencies which have started tracking drug related activities social media manually in the last year or so, it is also imperative to develop an automated solution.

Due to the inherent anonymity on Instagram, drug dealers always post their offers in the most blatant fashion, which provides us a good opportunity to track them through the contents they posted. For example, the Federal Government has used Instagram to arrest over 350 drug dealers and seize 7 million dollars in 2014 [5]. However, traditional approaches, such as manual search and browsing by trained domain experts, suffer from two major problems when dealing with social media data. (1) *Scalability*: it is impractical and inefficient to check the large-scale social media data manually. (2) *Reproducibility*: even domain experts may overlook the evidences of drug dealing in some cases due to human errors. Moreover, different people may have different criteria for drug dealers, which makes the result inconsistent and not reproducible.

To alleviate these problems, we propose to use machine learning algorithms to detect drug dealer accounts, which offers an effective and efficient way to assist the tracking of illicit drug trade. Some effort has been made recently to mine social media to track public health [Venkata Rama Kiran Garimella 2015], alcohol consumption [Pang et al. 2015] and even drug use patterns [Yiheng Zhou and Luo 2016] [Correia et al. 2016]. Nevertheless, detecting drug dealer accounts pose unique challenges. First, drug dealer accounts represent a very small population of over 300 million Instagram users. Second, the content of drug-related posts and accounts is diverse. For example, the image of a drug-related post can be pills, weeds or syrup in any environment (see sample images in Figure 4). Also, drug dealer accounts can be very different: some of them post most of drug-related posts while others do not. Third, domain knowledge is necessary in some ambiguous cases. For example, it is very difficult for a machine learning algorithm to differentiate a illicit drug dealer account from a legal one, or from a drug abuser account. These cases are common in our experiment and even trained experts may have different judgments.

In this paper, we propose a scheme to identify drug-related posts, analyze behavior patterns and finally detect drug dealer accounts. Multimodal data such as image, text, relational information and temporal pattern are used at different stages to improve the performance of the overall system. After training and testing on our sample datasets collected in the experiment, we can achieve over 88% accuracy on drug-related post recognition and 0.51 F_1 score on drug dealer detection (prior is about 1:8). Figure 1 provides an example of our approach in a real case, and also shows its effectiveness and efficiency: using only 1 hashtag ("xanx" in the example) as the search keyword and 30 search outcomes, our approach can effectively discover two drug dealer accounts ("*****237" and "*****420") on Instagram fully automatically.

This work makes the following contributions:

— We propose an effective multimodal approach to tracking drug dealing and abuse on Instagram.
— We propose a multi-task learning method to leverage web images for recognition tasks on social media.
— Our approach is generalizable to other similar problems, such as human trafficking and illegal gun sale.

[3] http://www.news4jax.com/news/local/teens-getting-drugs-through-instagram
[4] http://6abc.com/news/drug-dealers-use-social-media-to-target-kids/64262/
[5] http://wordondastreet.com/feds-use-instagram-arrest-350-drug-dealers-seize-7-million-one-weekend/

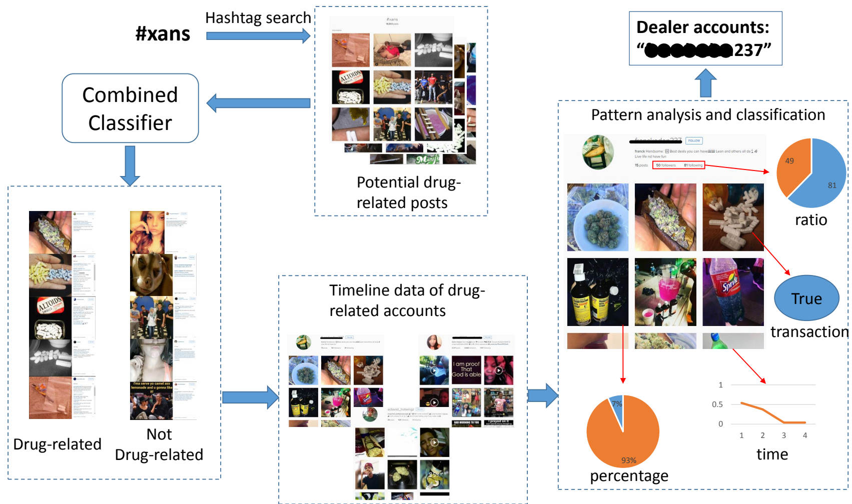


Fig. 1. Exemplary illustration of our proposed approach to detecting drug dealer accounts. “Percentage” indicates the percentage of drug-related posts (orange). “Ratio” indicates the ratio of the number of followers (orange) and followees (blue). “Time” indicates the frequencies of drug-related posts in hours of a day. “Transaction” indicates the evidence of transaction.

The organization of the remainder of this paper is as follows. In Sec. 2, we review the related work. In Sec. 3, we present our framework and data collection procedure. In Sec. 4, we describe the drug-related post recognition in detail, including the multi-task learning method for image-based classifier and the text-based classifier. Sec. 5 and Sec. 6 discuss the account pattern analysis and dealer account detection, respectively. Finally, Sec. 7 concludes the paper.

2. RELATED WORK

In this section, we briefly review two related research areas: social media data mining and multi-task learning.

2.1. Social Media Data Mining

Social media provides us a powerful new way of sensing social behaviors and activities from user-generated social multimedia contents. Many efforts have been made to mine social media data for various tasks, such as recommendation system [Yang et al. 2015], social affairs [Balasuriya et al. 2016] [Wang et al. 2015b] and public health [Venkata Rama Kiran Garimella 2015] [Pang et al. 2015]. Specifically, [Balasuriya et al. 2016] uses Twitter to find street gang members and achieves a promising F_1 score with classifiers trained on selected features. [Venkata Rama Kiran Garimella 2015] uses image posts on Instagram to predict different health-related variables. [Pang et al. 2015] monitors adolescent alcohol consumption using selfie photos and the associated tags on Instagram.

To the best of our knowledge, there is limited work on tracking drug abuse and dealing on social media. [Buntain and Golbeck 2015] proposes to analyze the time and location patterns of drug use by mining Twitter data. A more recent work, [Yiheng Zhou and Luo 2016], analyzes the drug use patterns using Instagram data. Drug-related posts are first determined by matching a dynamic hashtag set,time patterns and common interest patterns are then analyzed using data mining algorithms. [Correia et al. 2016] uses network information of Instagram user timelines to monitor potential drug interaction. While these works show great promise in using social media for drug use tracking, our work is differentiated from them primarily by proposing more robust drug-related post recognition methods and performing account-level analysis for dealer account detection.

2.2. Multi-task Learning

Multi-task learning is a representative approach to inductive transfer learning [Zheng 2015]. The idea of learning multiple tasks together allows the learner to leverage the information contained in the related tasks, which often leads to a better model for the main task. Many algorithms have been designed to solve multi-task learning problem, such as multi-task support vector machines [Evgeniou and Pontil 2004], multi-task feature selection [Liu et al. 2009] and multi-task structure learning [Argyriou et al. 2007]. Using a shared representation in neural networks is also a typical way to learn different tasks in parallel [Collobert and Weston 2008]. In [Wang et al. 2015a], the authors formalize the demographic prediction in retail scenario as a multi-task multi-class problem, then solve it by using structured prediction based on shared representation learning. Our multi-task learning approach for drug-related post recognition is inspired by [Wang et al. 2015a], but differs from it in some respects. For example, we introduce a mask mechanism and a task-relation encoding mechanism (details in Sec. 4.1) into the model, and our model has a simpler architecture than that in [Wang et al. 2015a].

3. FRAMEWORK OVERVIEW AND DATA COLLECTION

This section introduces our framework to identify drug-related posts, analyze behavior patterns and finally detect drug dealer accounts. We also introduce the data collection procedure and the datasets used for training our models.

3.1. Proposed Approach

Identifying illicit drug dealer accounts on Instagram is a challenging task, as the diversity of Instagram posts and accounts. Previous work, using hashtags to determine drug-related posts [Yiheng Zhou and Luo 2016], is not precise enough and produces many false-positive samples. To achieve better performance, we propose to divide the task into two stages: *post-level analysis* and *account-level analysis*. The first stage provides a robust model that can recognize a drug-related post with high accuracy, while the second stage provides a model to analyze account patterns and determine drug dealer account.

Specifically, our approach includes the following four steps. 1) *Potential posts collection*: Given a list of terms related to drug dealing, we collect a pool of potential drug-related posts by a hashtag-based search on Instagram. 2) *Drug-related post recognition*: Image-based and text-based classifiers are both trained and used to filter the potential posts. A multi-task learning method is employed to improve the image-based classifier using augmented datasets collected through a search engine. An intermediate decision is made by integrating the decisions of the two classifiers and corresponding drug-related accounts are identified. 3) *Account pattern analysis*: After collecting the timeline data of the drug-related accounts, we analyze various activity and behavior patterns of these accounts, including drug-related pattern, temporal pattern, relational information pattern, etc. 4) *Dealer accounts detection*: We determine drug dealer accounts using a pre-trained classifier based on a set of selected features. It is worthnoting that the feature selection procedure used in training the classifier can also help to improve the account pattern analysis stage. The entire procedure is shown in Figure 2.

Using the example in Figure 1, we further illustrate our approach to detecting drug dealer accounts in a real case. First, we randomly pick a term in the dictionary ("xans"

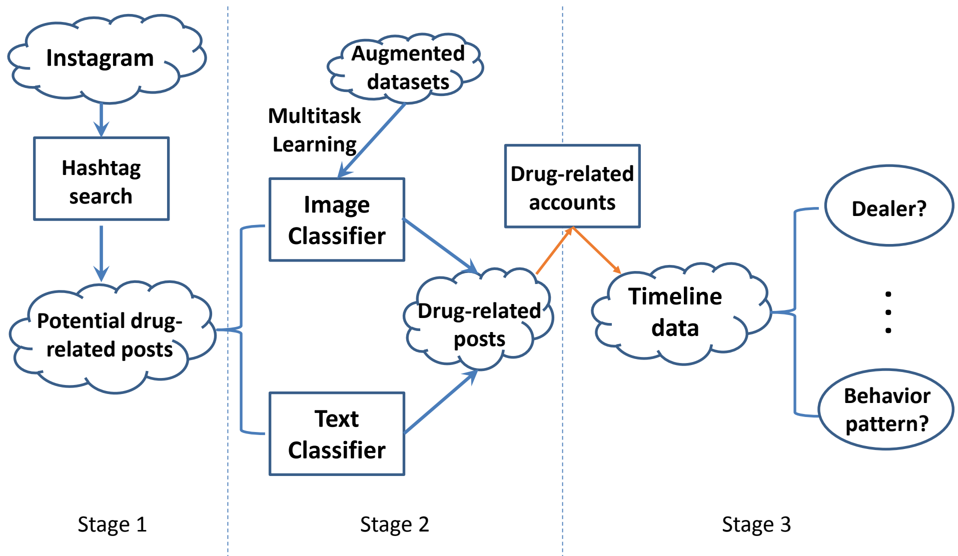


Fig. 2. The framework of our proposed approach.

in this example) and use it as the keyword to search on Instagram. For simplicity, we download the 30 most recent posts as an example. Second, using our trained image-based and text-based classifiers, we identify 18 of them as drug-related posts, including 11 unique accounts (5 posts of each class are shown). Finally, we collect the timeline data of these 11 drug-related accounts, extract their behavior patterns as features and apply our trained drug dealer classifier on them. As a result, two accounts ("*****237" and "*****420") are detected as drug dealer accounts and we prove the correctness by manually inspecting their pages.

3.2. Data Collection

We collect three sample datasets for our model training and testing. The *Instagram-post* and *Google-image* datasets are used in the post-level analysis stage, while the *Instagram-account* dataset is used in the account-level analysis stage. Table 3.2 summarize the statistics of our datasets.

Instagram-post dataset is collected through a hashtag-based search on Instagram. The search queries are obtained from a dictionary of terms related to drug dealing provided by domain experts (see Appendix). 4819 posts are collected and manually annotated as drug-related or not by experts. Each post includes image, hashtags and captions. Among these posts, only 1260 posts are annotated as positive, also indicating that hash-tag filtering alone is not robust enough.

Google-image dataset is collected as an augmented dataset for training the image-based classifier. There are two main reasons to build an augmented dataset. First, the appearances of drug-related posts are very diverse because of the different kinds of drugs and the intrinsic noise of social media posts. Second, the amount of labeled data

Table I. Summary of our datasets.

Dataset	# of samples	# of positive samples
Instagram-post	4819	1260
Instagram-account	206	27
Google-image	4329	675 / 2421 / 1233

Fig. 3. Sample images from Instagram-post dataset.

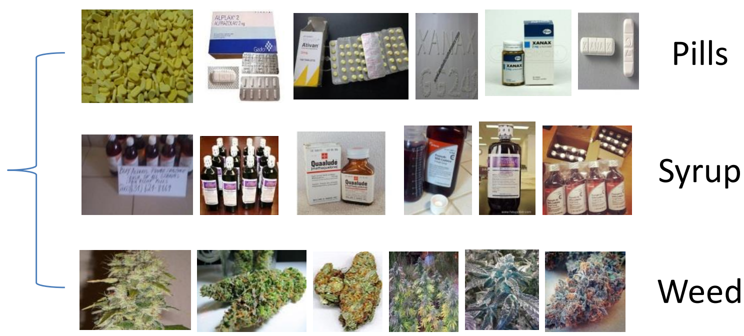


Fig. 4. Sample images from Google-image dataset.

in *Instagram-post* dataset is limited as manual annotation for images is tedious and time-consuming. As a result, it is difficult for a classifier trained on limited data to generalize well to diverse cases.

Web data provides us a good option to build this augmented dataset semi-automatically. Specifically, we collect data from an image search engine (Google Image Search) by searching the terms in the aforementioned dictionary as keywords. The outcomes can be viewed as data with noisy labels because they are representative images for the search keywords. The data is then cleaned and grouped into three predefined sub-classes based on their appearances: weed, pills and syrup. These three sub-classes cover the most common illicit drugs, as reported in [Yiheng Zhou and Luo 2016]. We collect 4329 images in total and group them into weed, pills and syrup, in the number of 675, 2421 and 1233, respectively. We also collect comparable amount of negative samples, which are obtained from data randomly collected on Instagram. Sample images from *Instagram-post* dataset and *Google-image* dataset are shown in Figure 3.

Instagram-account dataset is collected by crawling the timeline data of the drug-related accounts. In our experiment, we collect a sample dataset of 206 drug-related accounts and each of them has up to 200 recent posts. These accounts are picked according to the *Instagram-post* dataset: if a post is annotated as drug-related, the cor-

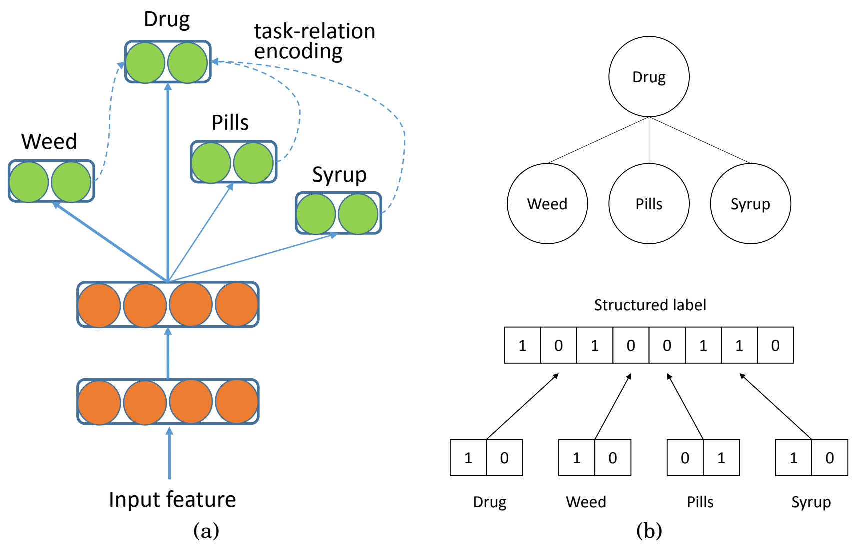


Fig. 5. (a) Model architecture for multi-task learning. Dash lines indicates "task-relation encoding". (b) Top: Task relationship can be encoded into a tree structure. Bottom: Structured representation of multi-task predictions.

responding account is viewed as a drug-related account. Each account in the dataset contains detailed information, including post content, temporal information of each post, the number of followers/followees, etc. These accounts are then annotated as drug dealer or not by two domain experts through inspecting their pages manually. Two sets of annotation are used to illustrate the problem of inconsistent human annotation mentioned in section 6.2. Finally, about 27 of these accounts are labeled as drug dealers with evidences of drug transactions.

4. DRUG-RELATED POST RECOGNITION

In this section, we describe how to train the classifiers for drug-related post recognition. Specifically, we first train the image-based and text-based classifiers separately. Decision-level fusion is then used to integrate two decisions as a weighted average.

4.1. Multi-Task Learning for Improved Image-based Classifier

Drug-related posts always contain images with drug-related content in them. As a result, the image-based classifier is essential to identify drug-related posts. As mentioned in section 3.2, data augmentation is necessary to build a robust classifier. Using image search engine-based data as representative images is also adopted in [Li et al. 2015; Chen and Gupta 2015]. However, it is nontrivial to combine the original and augmented datasets coherently for training. The first challenge is the different data domains: while one comes purely from Instagram, the other comes from diverse web data. The second challenge is the different annotation settings, which means the labels of different datasets have different meanings (drug, weed, pills, syrup). Therefore, we need to transfer the knowledge about both data and task from the source domain (web dataset) to the target domain (Instagram dataset).

We propose a simple multi-task learning method to solve this inductive transfer learning problem by learning common representations for relevant tasks [Collobert and Weston 2008]. Inspire by [Wang et al. 2015a], we first combine different label settings into a unified structured label representation, as illustrated in Figure 5 (b)

(Bottom). Moreover, task relationship is encoded into the structure representation. In our case, the four tasks can be represented as a tree structure as in Figure 5 (b) (Top), where a parent node is a super-concept of its children nodes. As a result, we activate the parent node (labeled as positive) if one of its children nodes is activated. We call it *task-relation encoding* for future reference. Intuitively, it means that whenever we see a “weed” image, we should also classify it as a drug-related image. The model architecture is based on a multi-layer neural network, in which the lower level layers are shared among different tasks and the top-most layers are independent for different predictions. We illustrate the architecture in Figure 5 (a). The objective function is

$$L = -\frac{1}{N}\sum_{n=1}^{N}\sum_{t\in T}\left[y_{nt}\log \hat{y}_{nt} + (1-y_{nt})\log(1-\hat{y}_{nt})\right] \tag{1}$$

where T indicates the set of tasks we consider for data n.

Several details are worth noting when optimizing the model. First, a masking mechanism is applied during training to ignore the irrelevant losses. Formally, as different nodes without direct connections should be mutually exclusive, we mask out the losses of all the unconnected nodes, given that the input data falls in a specific node. For example, when data of “weed” dataset is fed to the model, we ignore the losses calculated by the predictors of “pills” and “syrup”. In addition, although task-relation encoding introduces the dependency of tasks, this information is relatively noisy. Thus we set a lower weight (0.5) for the loss that comes from task-relation encoding prediction.

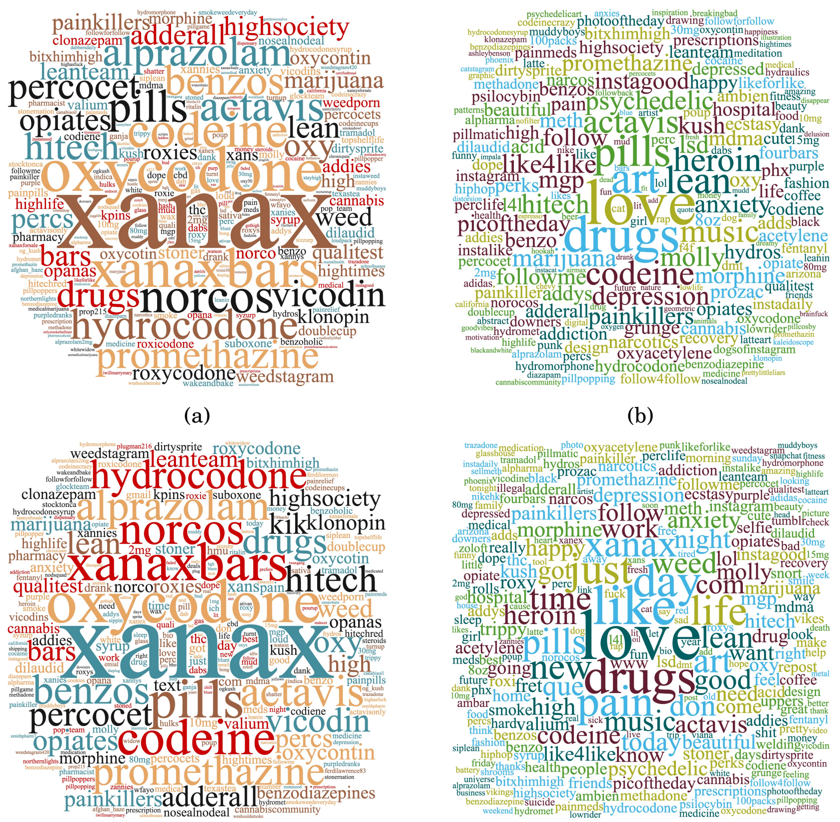
4.2. Text-based Classifier

Text data, including hashtags and captions, is also useful to differentiate drug-related and non-drug-related posts. Figure 6 shows the word clouds of hashtags and captions for two classes of posts in the *Instagram-post* dataset (words are lowercased and stemmed before processing, only the top 200 frequent words are drawn). For both hashtags and captions, we can see the clear difference between drug-related posts and non-drug-related posts. Specifically, it is much more frequent for drug-related posts to have words like “xanax”, “oxycodone” and “codeine”, while the non-drug-related posts contain the words like “love” and “day” more frequently. Note that all these posts in the dataset are collected by hashtag-based search, so the two classes share some common drug-related words such as “drug” and “actavis”. Another interesting finding from the figures is that, drug-related posts tend to be more concentrated on some specific words than non-drug-related posts, as the size of their frequent words are bigger than the others. This finding matches our common sense that the content of non-drug-related posts should be more diverse.

Next, we train a text-based classifier to recognize drug-related posts. We first extract different kinds of features for tags and captions. Specifically, we extract only uni-gram feature for tags and both uni-gram and bi-gram features for captions. Features are then scaled by tf-idf weighting [Salton and McGill 1986]. To reduce the feature dimension, we retain the top 1000 features ordered by term frequency across the corpus for tags and captions, respectively. As a result, we have a 2000 dimensional vector for each post. A naive Bayes classifier [McCallum et al. 1998] is used for text classification.

4.3. Decision Fusion and Evaluation

To better utilize multimodal data on social media, we use a decision-level fusion method to integrate the decisions of image-based and text-based classifiers. Another common method for multimodal fusion is to train a joint classifier, but it may not work very well for the social media data in our case. First, unlike typical highly correlated modalities such as video and audio, the correlation between image and text data is



(a) (b) (c) (d)

Fig. 6. Comparison of text used in the posts. First row: hashtags. Second row: captions. First column: drug-related posts. Second column: non-drug-related posts.

weaker on social media and often biased by the intrinsic noise. Therefore, training a joint model is usually challenging and ineffective. Second, training two separate classifiers can better capture unique information of different modalities, such as the diverse and subjective behaviors on social media. Finally, training a joint model would require much more paired training data of both modalities, which is not available in our case. Therefore, it is a sensible choice to use decision level fusion.

We evaluate the performance of post classification at this stage by training and testing on *Instagram-post* dataset. For the image-based classifier, we first extract visual feature of each image using GoogLeNet [Szegedy et al. 2015]. The features are then fed to the MTL framework as described above. We use two layers of shared hidden layers with the size of 256. Four softmax layers are connected for class prediction. The model is trained using gradient descent with RMSprop update rules. Both classifiers predict the probabilities of a post to be drug-related. A linear weighting method is used for decision-level fusion and we empirically pick the weight as 50% for the image-based classifier and 50% for the text-based classifier.

Five-fold cross validation is conducted to test the performance of the proposed methods. Different metrics: accuracy, precision, recall and f1 score are reported in Table II. The best performance is achieved by fusing two classifiers, with over 88% accuracy.

Table II. Results of drug-related post recognition.

Methods		Accuracy	Precision	Recall	F1-score
Image-based	MLP	86.9	80.3	66.3	0.72
	MT_MLP	87.2	81.0	66.7	0.73
Text-based	Tags only	82.2	67.2	62.0	0.65
	Caption only	81.7	69.4	53.7	0.61
	Combined	81.7	63.6	70	0.67
Late fusion		**88.1**	**83.1**	**68.1**	**0.75**

5. ACCOUNT PATTERN ANALYSIS

In this section, we analyze several behavior patterns of drug-related accounts using our *Instagram-account* dataset. We note that geolocation information is also useful for tracking drug dealers. However, as the location information of most posts are unavailable, we do not analyze the geolocation pattern of drug-related accounts.

5.1. Percentage of Drug-related Posts

With a trained classifier for drug-related posts, it is straightforward to obtain the percentage of drug-related posts within a user's timeline data, which is calculated by (# of drug-related posts / # of all posts). The value of the ratio can be used as a metric to distinguish drug dealer accounts on Instagram. Figure 7 shows the histogram of the percentages of dealer accounts and non-dealer accounts. It is clear that accounts with a higher value of drug-related percentage are more likely to be drug dealers, while most of the non-dealer accounts have a very low percentage of drug-related posts. Note that there are some drug dealer accounts have a very low value of the ratio (or vise versa), which shows the diversity of Instagram accounts.

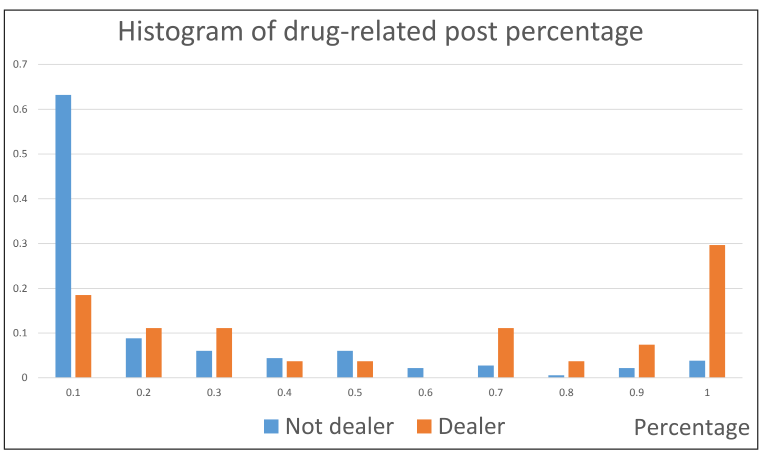


Fig. 7. Drug-related post ratio of dealers and non-dealers.

5.2. Temporal Patterns

Temporal patterns refer to the frequencies of drug-related posts along the time axis. We analyze the temporal patterns in different timescales, and find that the signal is most significant when we focus on different hours of a day. The statistics are shown in Figure 8. We can see that drug dealers tend to post more drug-related posts at midnight.

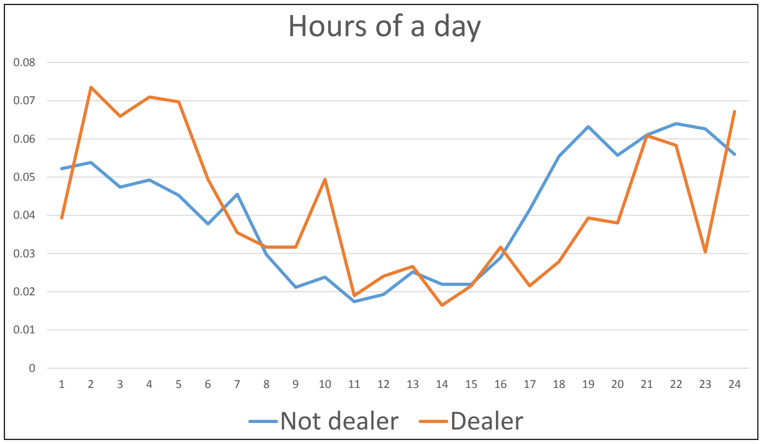


Fig. 8. Temporal pattern of drug-related accounts.

5.3. Relational Information

Relational information is also useful to distinguish drug dealer accounts. As expected, drug dealer accounts should have more followers than their following accounts, as they are selling their products to their followers. That means the ratio (# of following / # of follower) is quite low for them. Our analysis confirms this intuition in some sense, as shown in Figure 9.

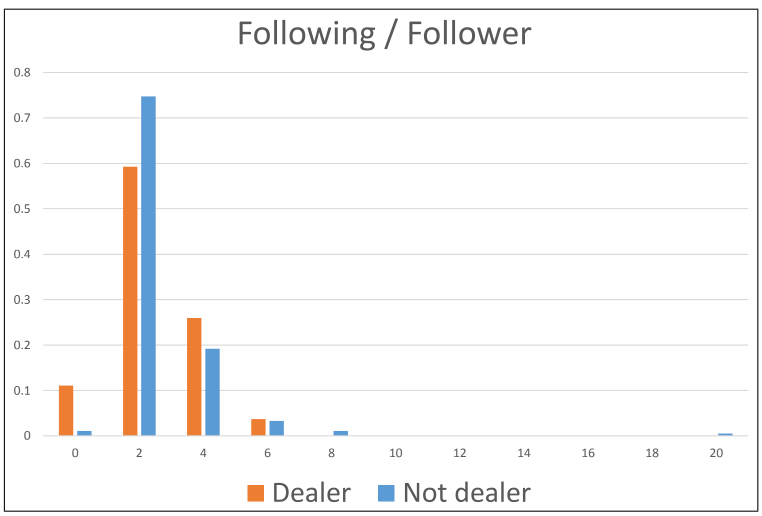


Fig. 9. Relational information pattern of drug-related accounts.

5.4. Evidences of Transactions

Evidences of transactions is another criterion to determine a drug dealer account. As stated by the domain experts, one account will not be identified as drug dealer account if it does not contain evidences of drug transactions, even though it contains many drug-related posts. This inspires us to capture the evidences of transactions in the timeline data. Specifically, we extract the bio description and comments of the drug-related posts. A predefined blacklist provided by the domain experts are applied to filtering the content. Once the terms of blacklist occur, we assign a 'True' value to this feature, otherwise we assign a 'False' value. Figure 10 shows the percentage of accounts with evidence of transactions. Even though our method cannot detect all the transaction information, we still can see the trend that dealer accounts have a higher percentage value than non-dealer accounts.

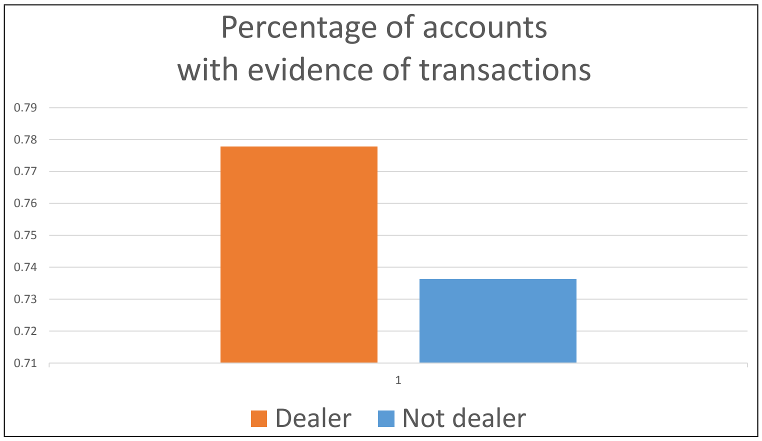


Fig. 10. Percentage of accounts with evidence of transactions.

6. DEALER ACCOUNT DETECTION

Finally, we use the aforementioned modalities with significant signals as features to detect drug dealer accounts. Specifically, we use the following features as the input:

— P: Percentage of drug-related post (1D)
— H_1 to H_4: Hour of a day (binned to 4D)
— R_1 to R_3: Relational information: # of Follower, # of Following, Following/Follower (3D)
— E: Evidence of transaction (1D)

Features are normalized to zero mean and unit variance.

6.1. Feature Selection

Although the above features show useful signals in pattern analysis, we are still not sure about their significance for detecting drug dealer accounts. We choose a feature selection approach using L1 regularization. In particular, we train a linear logistic regression classifier with L1 regularization on hold-out data and remove the feature dimensions with zero or very small coefficient values. Six features are retained after the feature selection, namely P, H_1 (midnight), H_4 (late night), R_1 (# of follower), R_3 and E. Five-fold cross validation is then performed to evaluate the method and the result is showed in Table III.

6.2. Inconsistency in Human Annotation

As we mentioned in Sec. 1, even domain experts may overlook some evidences or have inconsistent judgments during annotation. In this experiment, we have two independent sets of labels annotated by two experts. 94% of the labels are consistent between

Table III. Results of dealer account classification with two experts. We regard each expert's annotation as the ground truth to compute the evaluation scores for the other expert.

Method	Expert ID	Precision	Recall	F1-score
LR	exp1	28.8	55.6	0.38
LR	exp2	38.5	60.6	0.47
LR-select	exp1	30.6	55.6	0.40
LR-select	exp2	42.9	63.6	0.51

Table IV. Input and output of our classifier for three example accounts.

Account ID	*****728	*****230	*****976
P	0.93	0.12	0.31
H_1	0.54	0.75	0.29
H_4	0.042	0.042	0.21
R_1	49	273	43
R_3	1.65	3.40	0.14
E	1	0	1
Output	**1**	**0**	**0**
Ground truth	1	0	1

the two experts. However, such seemingly minor differences resulted in very different evaluation scores as our dataset is highly imbalanced (significantly fewer positives). Compared with annotation by human experts, our machine learning based approach offers reproducible and consistent prediction. Table III supports this observation.

6.3. Case study

In this section, we further discuss our approach by analyzing three example cases. Table 6.3 lists the values of five features and the output of our classifier for three accounts in our *Instagram-account* dataset. Ground truth is obtained by manually checking the pages of these accounts.

The first two cases show the effectiveness of our classifier for identifying drug dealer accounts. In particular, the features: percentage of drug-related posts (P), ratio of # of following and follower (R_3) and evidence of transaction (E) contribute to the correct classification and also match the analysis of behavior patterns in Sec. 5.

The third example is a false negative case. Although the features R_3 and E provide some supports for a positive prediction, the other features all vote for a negative output, especially the relatively low value of the feature P. This again shows the problem of diverse Instagram accounts and the complexity of the task. Our future work will focus on building more robust drug dealer account classifiers to alleviate this problem.

7. CONCLUSIONS

In this study, we propose a comprehensive framework for identifying drug-related posts, analyzing behavior patterns and detecting drug dealer accounts on Instagram. Multimodal data and analysis methods are employed to achieve high quality results competitive with domain experts. The experiments show that our proposed approach is effective and reproducible for practical use as a tool for tracking and combating illicit drug trade on social media. Our future work will focus on more robust account-level analysis: 1) collect larger account-level datasets semi-automatically using our current models, 2) investigate more useful features for robust classifiers, such as network information, and 3) analyze further account-level behavior patterns of drug dealers, such as their interaction patterns and grouping patterns.

APPENDIX

Terms Related to Illicit Drug Dealing

addys
Pills
benzodiazepines
8oz
hydromet
Painkillers
100packs
opiate
vicoden
trazadone
sellmeth
roxi
2mg
ecstasy
actavis
zannies
Percocet
MGP
xannys
10mg
Psilocybin
addies
Vicodin
Oxycodone
prescriptions
perclife
30mg
kpins
hydrocodone
promethazin
oxycontin
downers
ambien
percocets
benzodiazepine
diazapam
promethazineandcodeine
prozac
percocet
norocos
narcotics
vicodins
oxyacetylene
vicodine
zanned
adds
pillgame
klonazapem
muddyboys
Pillmatic
perc
xanies
fourbars
codeinekid
painkillers
narcos
tramadol
roxicodone
pills
molly
heroin
uppers
percosets
Magnacet
methadone
percs
MDMA
Trazadone
actavis
painmeds
15mg
ecstasy
promethazine
suboxone
phx
fentanyl
hydros
klonopin
xanex
benzos
roxy
Endocet
adderall
meth
vicodin
xans
pillpopping
xanax
oxycodone
alprazolam
Activis
alpharma
mdma
opiates
painkiller
xannies
xanax4sale
morphine
nosealnodeal
alproloz
xanexfootballs
roxies
clonazepem
vikes
Zoloft
promethazinecodeine
alprazolam2mg
acetylene
clonazepam
oxy
roxycodone
80mg
xannies
xanaxbars
codiene
anxies
hydromorphone
benzo
musclerrelaxers
codeine
snort
perks
valium
dilaudid
rolls
az
oxycodone
klonazepam
hydrocodonesyrup
roxys
oxycoti

ACKNOWLEDGMENTS

We thank the support of New York State through the Goergen Institute for Data Science, and the encouragement and assistance by Lacey Kelly and Meredith McCarron of the Office of the New York State Attorney General.

REFERENCES

Andreas Argyriou, Charles A. Micchelli, Massimiliano Pontil, and Yiming Ying. 2007. A Spectral Regularization Framework for Multi-Task Structure Learning. *Advances in Neural Information Processing Systems* 20 (2007), 25–32.

Lakshika Balasuriya, Sanjaya Wijeratne, Derek Doran, and Amit Sheth. 2016. Finding Street Gang Members on Twitter. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2016)*, Vol. 8.

Cody Buntain and Jennifer Golbeck. 2015. This is your Twitter on drugs: Any questions?. In *International Conference on World Wide Web*.

Xinlei Chen and Abhinav Gupta. 2015. Webly supervised learning of convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision*. 1431–1439.

Ronan Collobert and Jason Weston. 2008. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning*. ACM, 160–167.

Rion Brattig Correia, Lang Li, and Luis M Rocha. 2016. Monitoring Potential Drug Interactions and Reactions via Network Analysis of Instagram User Timelines. In *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, Vol. 21. NIH Public Access, 492.

Theodoros Evgeniou and Massimiliano Pontil. 2004. Regularized multi–task learning. In *Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August*. 109–117.

Yuncheng Li, Xitong Yang, and Jiebo Luo. 2015. Semantic Video Entity Linking Based on Visual Content and Metadata. In *Proceedings of the IEEE International Conference on Computer Vision*. 4615–4623.

Jun Liu, Shuiwang Ji, and Jieping Ye. 2009. MultiTask Feature Learning Via Efficient $\ell_{2,1}$Norm Minimization. *Eprint Arxiv* (2009), 339–348.

Andrew McCallum, Kamal Nigam, and others. 1998. A comparison of event models for naive bayes text classification. In *AAAI-98 workshop on learning for text categorization*, Vol. 752. Citeseer, 41–48.

Ran Pang, Agustin Baretto, Henry Kautz, and Jiebo Luo. 2015. Monitoring adolescent alcohol use via multimodal analysis in social multimedia. In *Big Data (Big Data), IEEE International Conference on*. 1509–1518.

Gerard Salton and Michael J. McGill. 1986. *Introduction to Modern Information Retrieval*. McGraw-Hill, Inc., New York, NY.

Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2015. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1–9.

Ingmar Weberr Venkata Rama Kiran Garimella, Abdulrahman Alfayad. 2015. Social Media Image Analysis for Public Health. *arXiv preprint arXiv:1512.04476* (2015).

Pengfei Wang, Jiafeng Guo, Yanyan Lan, Jun Xu, and Xueqi Cheng. 2015a. Your Cart tells You: Inferring Demographic Attributes from Purchase Data. (2015).

Yu Wang, Jianbo Yuan, and Jiebo Luo. 2015b. America Tweets China: A fine-grained analysis of the state and individual characteristics regarding attitudes towards China. In *IEEE International Conference on Big Data*. 936–943.

Xitong Yang, Yuncheng Li, and Jiebo Luo. 2015. Pinterest Board Recommendation for Twitter Users. In *Proceedings of ACM Conference on Multimedia*. 963–966.

Numair Sani Yiheng Zhou and Jiebo Luo. 2016. Understanding Drug Use Behaviors by Mining Instagram. In *International Conference on Social Computing, Behavioral-Cultural Modeling & Prediction and Behavior Representation in Modeling and Simulation (SBP-BRiMS)*.

Yu Zheng. 2015. Methodologies for Cross-Domain Data Fusion: An Overview. *Big Data IEEE Transactions on* 1 (2015), 16–34.