Inferring Social Influence of Anti-Tobacco Mass Media Campaign

Qianyi Zhan,* Jiawei Zhang, Philip S. Yu, Sherry Emery, and Junyuan Xie

***Abstract*— Anti-tobacco mass media campaigns are designed to influence tobacco users. It has been proved that campaigns will produce users' changes in awareness, knowledge, and attitudes, and also produce meaningful behavior change of audience. Anti-smoking television advertising is the most important part in the campaign. Meanwhile, nowadays, successful online social networks are creating new media environment, however, little is known about the relation between social conversations and anti-tobacco campaigns. This paper aims to infer social influence of these campaigns, and the problem is formally referred to as the Social Influence inference of anti-Tobacco mass mEdia campaigns (Site) problem. To address the Site problem, a novel influence inference framework, TV advertising social influence estimation (Asie), is proposed based on our analysis of two real anti-tobacco campaigns. Asie divides audience attitudes toward TV ads into three distinct stages: 1) cognitive; 2) affective; and 3) conative. Audience online reactions at each of these three stages are depicted by Asie with specific probabilistic models based on the synergistic influences from both online social friends and offline TV ads. Extensive experiments demonstrate the effectiveness of Asie.**

***Index Terms*— TV advertising, anti-tobacco, social network analysis, public health.**

I. Introduction

SMOKING remains the leading cause of preventable death and disease in the United States, killing more than 480,000 Americans each year (CDC, 2015). Anti-tobacco mass media campaigns, which are conducted to build public awareness of the immediate health damage caused by smoking and encourage smokers to quit, continue to represent an important strategy employed by the public health community to address this concern. Their influence on smoking behavior has been comprehensively studied. Strong evidence links media campaigns with reduced smoking among adults and youth, more favorable attitudes toward tobacco control, and increases in cessation behavior [34] and [13]. Campaigns usually use an ad placement strategy, including placing ads in print publications, outdoor venues, television, radio and social media.

Anti-tobacco campaigns succeed by placing well-crafted messages where their target audiences are likely to see or hear them. Viewers increasingly use multiple screens to simultaneously interact with different media platforms. According to report from Ericsson ConsumerLab, 62% of people worldwide reported using social networking sites and forums while watching television [14]. In June 2012 one in three Twitter users reported posting tweets about television content while viewing, an increase of 27% from only five months prior [14]. YouTube users actively engage with television programs on subscribed channels, and 90% of TV viewers visit YouTube or Google Search to extend their experiences with favorite programs [8]. Social media can increase the amount of exposure to the campaign and amplify the effect of TV exposures to gain a larger audience. Moreover, they can provide campaigns with important feedback on perceived effectiveness of a TV ad, and also reinforce the viewersâŁ™ engagement with campaign content. However little is known about how anti-tobacco campaigns are related to the social media conversation, and what extent the social conversation stimulates further engagement with the campaign.

Motivated by this, in this paper, we will learn the information propagation process from two anti-tobacco mass media campaigns: "CDC Tips" and "Legacy Truth", and understand different roles played by traditional (TV advertising) and social conversation (Tweets) in each campaign. This problem is proposed as the "Social Influence inference of anti-Tabcco mass mEdia campaigns (Site)" problem.

The Site problem is a novel problem, and it is very different from existing works on TV advertising studies in various disciplines, such as *social science* [35], *marketing* [5] and *advertising* [29]. There are also some works about Social TV in *human-computer interaction* area. For example, [32] explores motivations for live-tweeting across a season of a television show. Their work (1) aims to discover the motivation of tweeting, which is different from prediction in our problem and (2) is conducted on a TV series, thus the conclusion cannot applied on TV advertising. [9] developed a model for predicting TV audience rating according to word-of-mouse on

Manuscript received May 5, 2017; accepted May 12, 2017. Date of publication May 23, 2017; date of current version August 11, 2017. This work was supported in part by NSF under Grant IIS-1526499, Grant CNS-1626432, and Grant NSFC 61672313, in part by several awards: NCI/NIH and FDA Center for Tobacco Products under Award P50CA179546, in part by CDC under Award U01CA154254-05S1, in part by research grant from the Truth Initiative Foundation, in part by the National Key R&D Program of China under Grant 2016YFB1001102, and in part NSFC under Grant 61375069, Grant 61403156, and Grant 61502227. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NCI, NIH, FDA, CDC, or Truth. *Asterisk indicates corresponding author.*

*Q. Zhan is with the National Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China (e-mail: zhanqianyi@gmail.com).

J. Zhang and P. S. Yu are with the University of Illinois at Chicago, Chicago, IL 60607 USA (e-mail: jzhan9@uic.edu; psyu@uic.edu).

S. Emery is with NORC, The University of Chicago, Chicago, IL 60603 USA (e-mail: emery-sherry@norc.org).

J. Xie is with the National Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China (e-mail: jyxie@nju.edu.cn).

Digital Object Identifier 10.1109/TNB.2017.2707075

Facebook, which is opposite to the object of the SITE problem. Distinct from these existing works, this paper is the first to study the relation between TV advertising and social activities in data mining area.

Besides its importance and novelty, the SITE problem is very challenging to solve due to the following reasons:

- *Audience Attitude Modeling*: The effects of TV advertising programs on different audience will be very different. Even the same TV ads can usually evoke very diverse reactions of the audience, e.g., some audience like the ads but some don't. An effective modeling of the audience attitudes toward the TV advertising programs can be the prerequisite for inferring the potential social activities of the audience regarding the ads.
- *Synergistic Influence from Multiple Sources*: audience in online social networks can receive information about the TV ads from multiple sources, including both offline TV advertising programs and audience online friends. A new diffusion model which can effectively fuses the synergistic effects of these diverse influence sources on audience is needed.

To resolve these two challenges in the SITE problem, a new TV ads influence inference framework, "TV *Advertisements Social Influence Estimation*" (ASIE), is introduced in this paper. From the perspective of psychology [4], [10], [16], ASIE divides audience reactions and attitudes toward TV ads into three distinct stages: (1) Cognitive, (2) Affective and (3) Conative. Furthermore, ASIE depicts online audience actions on each stage with 3 specific probabilistic models. These synergistic influence from both offline TV ads and audience online friends is effectively fused in ASIE with the Poisson binomial distribution. Various parameters involved in the probabilistic distribution models can be learned automatically from historical data in ASIE.

The rest of paper is organized as follows: Section 2 describes the two real-world TV advertising campaign datasets. The problem will be formally defined in Section 3. In Section 4, the ASIE model will be introduced in details, whose performance is evaluated in Section 5. Finally, we discuss the related works in Section 6 and conclude this paper in Section 7.

II. ANTI-TOBACCO MASS MEDIA CAMPAIGNS

An anti-tobacco campaign refers to a series of ads programs that are broadcast through different media channels within a specific time range to build public awareness of the immediate health damage caused by smoking and encourage smokers to quit. Campaigns usually combine online channels, e.g., online social media, and offline channels, like the TV broadcasting, radio and print publications, which TV ad is the most important part among them. In this paper we evaluate two anti-tobacco campaigns upon data collected from both the TV ads records and the Twitter social network.

A. Data Analysis Settings

The TV broadcasting information is provided by the agency which conduct the campaign. Each ad record in the TV dataset contains its broadcasting time and its Nielson rating. Nielsen ratings are the audience measurement systems to determine the TV audience size and one single national ratings point represents 1% of the total number, or 1, 156, 000 households for the 2013-14 season [1]. The Twitter posts related to the advertising campaign are collected by using a large number of correlated keywords and hashtags by the company GNIP,[1] which provides data from dozens of social media websites via APIs. After getting the raw data from GNIP, we cleaned the data manually to remove the irrelevant tweets. The authors of crawled tweets are regarded as infected users, whose social connections are further crawled with the public API provided by Twitter.

Fig. 1. "Time window" example.

To measure the relationship between the number of related tweets and the corresponding TV ad ratings, different correlation metrics are applied, including both the Pearson and Spearman correlation coefficients. In statistics, the *Pearson Correlation Coefficient (PPMC)* [15] measures the linear correlation between two variables, giving a value between $+1$ and -1 inclusive, where 1 is total positive correlation, 0 is no correlation, and -1 is total negative correlation. While the *Spearman rank correlation coefficient* [11] assesses how well the relationship between two variables can be described using a monotonic function. If there are no repeated data values, a perfect Spearman correlation of $+1$ or -1 occurs when each of the variables is a perfect monotone function of the other.

Let the *time window* d (with length t_d) denote the time range before a tweet is posted. For example, as Fig. 1 shows, if tweet b is posted by user u at 2 pm, and we set $t_d = 1$ hour, its time window will be $d_b = [1\ pm, 2\ pm]$. While $t_d = infinte$ means we trace back to the start time of the entire campaign. The set of TV ads aired within the time window is represented as S_b^{tv}, and the tweets set is S_b^{sn}, which includes all tweets posted by u's social friends. In addition, exposures set is denoted as $S_b = S_b^{tv} \cup S_b^{sn}$, where $S_b = \emptyset$ implies u receives no exposures at all. In the example, $S_b^{tv} = \{$TV ad 2, TV ad 3$\}$, $S_b^{sn} = \{$tweet a$\}$ and $S_b = \{$TV ad 2, TV ad 3, tweet a$\}$. We gather all the tweets whose exposure set is not empty, i.e. $M_{t_d} = \{b | S_b \neq \emptyset\}$ and calculate its proportion among all the crawled tweets (of size N), i.e. $P_{t_d} = \frac{|M_{t_d}|}{N}$, where N is the number of all tweets. This proportion indicates the percent of users who have the chance to get information about campaigns during t_d. Similarly, we can get the ratios $P_{t_d}^{tv}$ and $P_{t_d}^{sn}$ to denote the proportion of users who receive the exposures from TV and online friends respectively.

Now we can analyze the two anti-tobacco campaigns based on the above measurements.

B. CDC Tips

The first advertising campaign is "Tips from Former Smokers 2013", launched by Centers for Disease Control and

[1] https://gnip.com

TABLE I
TWITTER STATISTIC SUMMARY

	CDC Tips	Legacy Truth
Date	Mar. 1 - Jun. 23	Aug. 1 - Oct. 31
Twitter	146,759	59,605
Retweets	46,402	45,676
Users	126,327	47,852
Tweets per Users	1.162	1.246
Edges	76,916	30,275
Followers Median	331	480
Followers Max	2,853,320	14,857,309

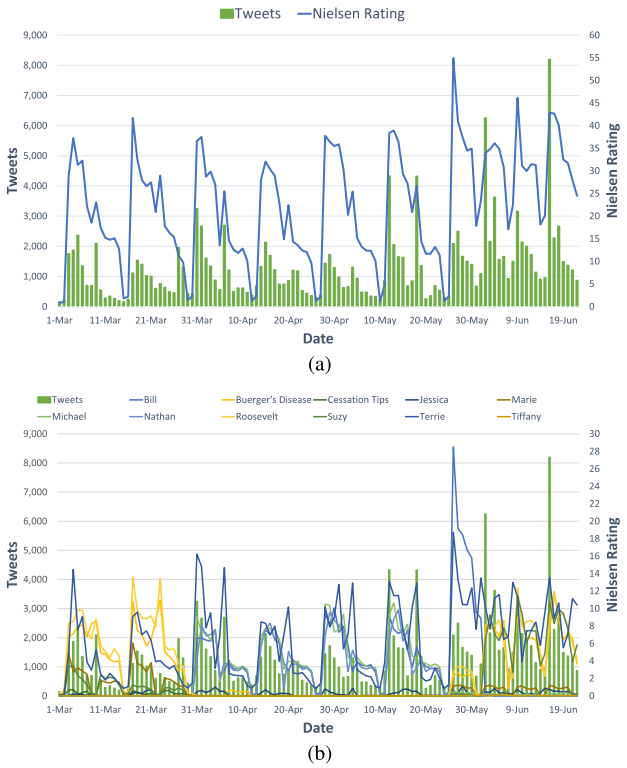

Fig. 2. Correlations between tweets amount and TV rating of "CDC Tips". (a) All. (b) Different topics.

Prevention (CDC), and it is hereinafter referred as the "CDC Tips" for simplicity. The "CDC Tips" was the federal government's first nationwide effort to use paid advertising to promote smoking cessation. "CDC Tips" advertising campaign began on March 1st and ended at June 23rd in 2013, which contained 10 different stories from 10 former smokers.

We crawled tweets related to the "CDC Tips", and their authors' profile from Twitter. The basic statistical information is available in the "CDC Tips" column of Table I. In summary, this campaign generated a total of 146,759 tweets related to the televised ads, i.e., 1,277 tweets per day on average.

1) Is Audience Reaction in Twitter Correlated With the TV Ratings?: Fig. 2 shows the number of tweets and TV ratings for the entire campaign and different stories. Both the Pearson correlation coefficient (0.64) and Spearman rank correlation (0.83) report a strong positive relationship between ratings and tweets in Fig. 2(a). As shown in Fig. 2(b), among all stories of "CDC Tips", *"Terrie"* exhibited the strongest correlation in both the Pearson correlation coefficient (0.64) and Spearman rank correlation (0.80).

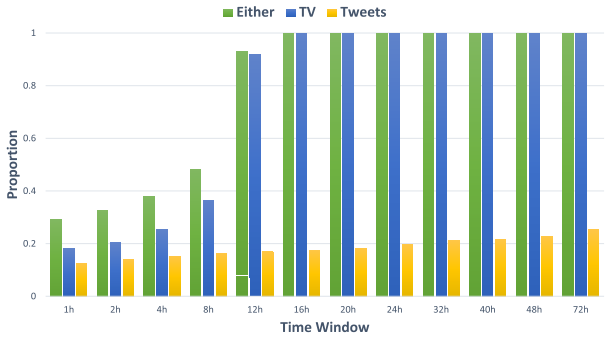

Fig. 3. Proportion of users who can get exposures with different time windows of "CDC Tips".

2) Does Audience React Immediately After Being Exposed to TV Ads and Tweets?: We change the length of time window t_d and calculate the user proportion of who can get information from TV, Twitter and either way, which can be represented as the ratios $P^{tv}_{t_d}$, $P^{sn}_{t_d}$ and P_{t_d} respectively. The statistical results with different time windows are shown in Fig. 3, which counters the intuition that people will tweet as soon as they see these exposures. When the t_d is 1 hour, more than 70% of the users cannot receive any kind of exposure, i.e., $S_b = \emptyset$. Until t_d is extended to 12 hours, the majority (93.2%) of the users can get campaign messages, but mostly from the offline TV ads. The information obtained from the online social network is very limited, and even tracing backward for 3 days, only one quarter (25.4%) of the users can receive exposures from their social friends. This may be the result of that "CDC Tips" did not do much online marketing in the Twitter network.

C. Legacy Truth

The other advertising campaign, "Legacy Truth", is launched by American Legacy Foundation (Legacy), which is national public health organization devoted to tobacco-use prevention. "Legacy Truth" provides young people with various concrete examples about the hazards of smoking, and encourages the teens generation to have a healthy lifestyle free from the smoking habits. "Legacy Truth" is actually a year-round advertising campaign, and we only take one segment of the campaign during August 11 and October 28 in 2013. "Legacy Truth" paid the majority of their attentions on traditional TV advertising, and have also broadcast their ads during the 2013 MTV Video Music Awards. Meanwhile it also initiated the explorations of disseminating their campaign information through online social media by promoting specific hashtags, and inviting some celebrities to join in the activities.

As shown in Table I, 59,605 tweets correlated to the campaign are crawled during August 11 - October 28, 2013. On average, 647.88 tweets are posted on each day. Significantly, 76.6% of the tweets are generated by retweeting. Similar to the "CDC Tips" part, we also analyze the "Legacy Truth" dataset from the following two directions:

1) Is Audience Reaction in Twitter Correlated With the TV Rating?: For the "Legacy Truth" dataset, we also measure the correlation between TV ratings and the number of tweets based on the Pearson and Spearman correlation coefficients

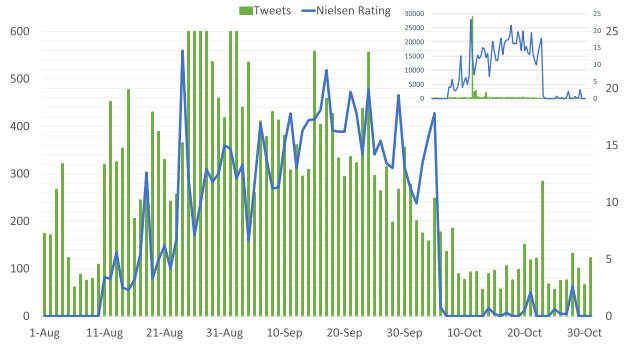

Fig. 4. Correlations between tweets amount and TV rating of "Legacy truth".

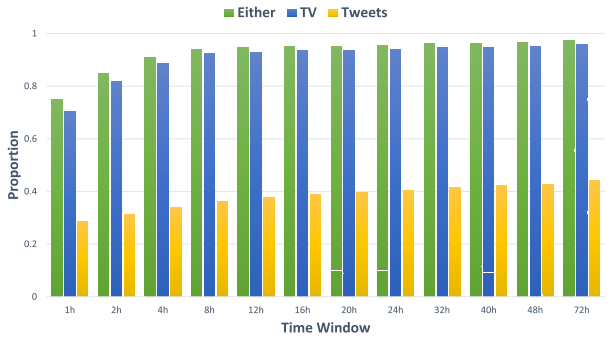

Fig. 5. Proportion of users who can get exposures with different time windows of "Legacy truth".

respectively. The result is shown in the small plot at the upper right corner of Fig. 4, where TV rating reached a high peak on Aug, 24 since "Legacy Truth" ads were aired during 2013 MTV Video Music Awards. Moreover, the number of tweets post also rose dramatically and reached the peak at 28, 958 on August 25 as some music stars were also discussing about "#Truth" in Twitter at the same time. The Pearson correlation coefficient (0.48) and Spearman rank correlation (0.79) demonstrate that the TV ratings and tweets amount have strong correlation. The peak point in data makes the relation in other normal days not obvious. Therefore, we set the limit of tweets number axis as 600, and as shown in the main figure of Fig. 4, the strong correlation between TV rating and tweets can be observed.

2) Does Audience React Immediately After Being Exposed to TV Broadcasting and Tweets?: Fig. 5 shows the proportion of tweets' authors who can be influenced by TV and social media exposures with different time windows. Unlike "CDC Tips", most users (75.1%) can receive campaign message in 1 hour in the "Legacy Truth", because of its intense TV advertising. Moreover, the ratio of users to receive exposures from their social friends is also high, since this campaign made an effort on viral marketing. However, it is interesting that even when t_d is 3, two ways of spreading information still cannot cover all users, which means a small part of users (2.5%) get information through other channels.

III. Problem Formulation

After analyzing two anti-tobacco campaigns preliminary, we propose our model to further study how TV ads affects social conversations. In this section, we begin with some important concepts mentioned in the paper and then formulate the problem.

One *advertising campaign* $\mathcal{C}$ is a series of ad messages that share the same idea which make up an integrated marketing communication. It usually has several creative themes (topics), which are formally presented as $t_1, t_2, \ldots, t_d$, where d is the number of the topics. As the analysis above, advertising campaigns mainly utilize both offline public media (TV advertising) and online social media to help influence more.

In TV advertising, ads are televised repeatedly to get the attention of a potential customer and each repetition is regarded as one *TV appearance* and all of them comprise a *TV stream*.

Definition 1 (TV Appearance): A TV appearance a is defined as a vector $a = (\tau_a, t_a, r_a)$, where τ_a is a's displayed time, $t_a \in A$ denotes the ad topic of a and r_a represents a's TV rating.

Since audience will not wait to see the ads, TV rating r_a is only correlated with a's displayed time τ_a. However different topics can evoke audience different reactions.

Definition 2 (TV Stream): The TV stream is the list of TV appearances $\mathcal{S}^{tv} = (a_1, a_2, \ldots, a_m)$, where m is the size of TV stream.

To formalize *social media* conversation, we first model the Twitter network. In the social network, each post related to the campaign $\mathcal{C}$, is regarded as a *social network (SN) appearance*.

Definition 3 (Online Social Network (OSN)): An online social network is a graph $G = (V, E)$, where a node $u \in V$ represents a user, and an edge $e = (u, v) \in E$ represents a connection between the users u and v. In our case, edges are all directed.

Definition 4 (Social Network (SN) Appearance): A SN appearance is defined as a post related to $\mathcal{C}$, denoted as $b = (\tau_b, t_b, w_b)$, where τ_b is the posting time of b, t_b denotes the b's topic and $w_b \in V$ represents the author, who is activated at t_b.

All of related *SN appearances* compose *SN stream* of the advertising campaign.

Definition 5 (SN Stream): The SN stream is the list of related SN appearances $\mathcal{S}^{sn} = (b_1, b_2, \ldots, b_n)$, where n is the size of the SN stream.

Based on the above definitions, we formally define the problem of *Social Influence Inference of TV Advertising*.

*Definition 6 (Social Influence Inference of TV Advertising (*ASIE*) Problem):* Given an advertising campaign $\mathcal{C}$, with its TV stream $\mathcal{S}^{tv}$ and the SN stream $\mathcal{S}^{sn}$ based on the social network $G = (V, E)$, the aim of SITE problem is to predict the probability of any user u being activated by campaign $\mathcal{C}$ at time τ, i.e., $P_u(\tau)$.

IV. Model Framework

In this section, we develop a novel model: TV Advertising Social Influence Estimation (ASIE) Model, to estimate social influence caused by TV ads and predict users' activation probability by incorporating the TV ads effect and friends influence in online social networks.

TABLE II
NOTATIONS AND DESCRIPTION AT THREE STAGES

Kind	Stage	Description	Probability	Distribution	Deciding Factors
TV appearance a	Cognitive	u remembers a at time τ.	$P(h_{u,\tau}^{tv,a}=1)$	Bernoulli $(\alpha_{u,\tau}^{tv,a})$	r_a, $(\tau-\tau_a)$
	Affective	u is impressed by a.	$P(l_u^{tv,a}=1)$	Bernoulli$(\beta_u^{tv,a})$	Normal distribution
	Conative	u is influenced by a at time τ.	$P(g_{u,\tau}^{tv,a}=1)$	Bernoulli $(\gamma_{u,\tau}^{tv,a})$	$\alpha_{u,\tau}^{tv,a}\times\beta_u^{tv,a}$
SN appearance b	Cognitive	u remembers b at time τ.	$P(h_{u,\tau}^{sn,b}=1)$	Bernoulli $(\alpha_{u,\tau}^{sn,b})$	$\omega(w_b,u)$, $(\tau-\tau_b)$
	Affective	u is impressed by b.	$P(l_u^{sn,b}=1)$	Bernoulli$(\beta_u^{sn,b})$	Normal distribution
	Conative	u is influenced by b at time τ.	$P(g_{u,\tau}^{sn,b}=1)$	Bernoulli $(\gamma_{u,\tau}^{sn,b})$	$\alpha_{u,\tau}^{sn,b}\times\beta_u^{sn,b}$
All	Aggregation	u is activated by k appearances at τ.	$F_{u,\tau}(k)$	Poisson binomial	

As we noted above, existing information diffusion models which take external events into consideration are proposed for news and popular social trends. However these models cannot be applied directly on TV advertising, because the aim of advertising is different and user feelings evoked by advertising is more complicated. Since the consumer attitude has been extensively researched in psychology and marketing area, our ASIE model is designed based on classical model mentioned in both social psychology [16] and marketing theories [4], [10]. The theory defines that user attitude has three stages: *Cognitive*, *Affective* and *Conative*. We modify these stages to fit our case and explain them in detail as following.

Definition 7 (Cognitive): At first audience become knowledge aware. In the ASIE model, this stage represents that users gather knowledge from TV and SN appearances.

Definition 8 (Affective): This stage ensures audience having strong feelings on the advertising. In the ASIE model, affective means TV and SN appearances is impressive to users.

Definition 9 (Conative): On this stage, audience have tendency to take action toward $\mathcal{C}$. In the viral marketing, the action is defined as posting tweets related to $\mathcal{C}$.

User attitudes toward a TV appearance a and SN appearance b are different, thus we discuss them separately and aggregate all appearances at last. Notations and description at three stages are listed in Table II.

A. TV Appearances

We first focus on user attitude of three stages toward a specific TV appearance a.

1) Cognitive: At the cognitive stage, $h_{u,\tau}^{tv,a}$ denotes at time τ whether user u remembers TV appearance a, formally defined as

$$h_{u,\tau}^{tv,a}=\begin{cases}1 & \text{if } u \text{ remembers } a \text{ at } \tau\\ 0 & \text{otherwise}\end{cases}$$

Since $h_{u,\tau}^{tv,a}$ is a binary valued variable drawn from Bernoulli distribution with mean $\alpha_{u,\tau}^{tv,a}$ [26], i.e.

$$h_{u,\tau}^{tv,a}\sim Bernoulli(\alpha_{u,\tau}^{tv,a})$$

The value of $\alpha_{u,\tau}^{tv,a}$ depends on the factors which affect the probability of a TV appearance being remembered. An intuitive thinking is the time lapse $(\tau-\tau_a)$, since time effaces memory. The longer time interval is, the less impression the appearance leaves. Therefore the value of $\alpha_{u,\tau}^{tv,a}$ is negative correlated to $(\tau-\tau_a)$. While as to a TV appearance, another related factor is TV rating, which indicates its audience size. High TV rating implies more audiences have watched it and the probability of an individual receiving information is higher. So the value of $\alpha_u^{tv,a}$ is positive related to r_a. Thus we choose the following function to calculate $\alpha_{u,\tau}^{tv,a}$.

$$\alpha_{u,\tau}^{tv,a}=r_a\times\theta^{tv}e^{-\theta^{tv}(\tau-\tau_a)} \tag{1}$$

Here the exponential function $e^{-\theta^{tv}(\tau-\tau_a)}$ is used to describe the decay of time effect, and θ^{tv} is the parameter will be learned in next part. We normalize the value of r_a in $[0,1]$ to make sure $\alpha_{u,\tau}^{tv,a}\in[0,1]$.

2) Affective: User emotion to this TV appearance is evoked at the affective stage. Similar to the stage of cognitive, $l_u^{tv,a}$ represents whether u is touched by TV appearance a, and is also drawn from a Bernoulli distribution with the mean $\beta_u^{tv,a}$.

$$l_u^{tv,a}=\begin{cases}1 & \text{if } u \text{ is impressed by } a\\ 0 & \text{otherwise}\end{cases}$$
$$l_u^{tv,a}\sim Bernoulli(\beta_u^{tv,a})$$

The value of $\beta_u^{tv,a}$ indicates the probability of u is impressed by a. It varies with individuals, and we assume it is randomly sampled from a normal distribution, with parameters μ_{t_a} and σ_{t_a} decided by a's topic t_a.

$$\beta_u^{tv,a}=\frac{1}{\sigma_{t_a}\sqrt{2\pi}}e^{-\frac{(x-\mu_{t_a})^2}{2\sigma_{t_a}^2}} \tag{2}$$

3) Conative: If u remembers a and u is impressed by a at the same time, u maybe intends to discuss $\mathcal{C}$ in social networks, which we say u is influenced by a, represented as:

$$g_{u,\tau}^{tv,a}=\begin{cases}1 & \text{if } u \text{ is influenced by } a \text{ at } \tau\\ 0 & \text{otherwise}\end{cases}$$
$$g_u^{tv,a}\sim Bernoulli(\gamma_{u,\tau}^{tv,a})$$

Based on our assumption, we define the *influence probability*.

Definition 10 (Influence Probability): The probability of u being influenced by TV appearance a is

$$P(u \text{ is influenced by } a \text{ at } \tau)$$
$$= P(u \text{ remembers } a \text{ at } \tau)\times P(u \text{ is impressed by } a)$$

From the definition, we get the value of $\gamma_{u,\tau}^{tv,a}$:

$$\begin{aligned}\gamma_{u,\tau}^{tv,a} &= P(g_{u,\tau}^{tv,a}=1)\\ &= P(h_{u,\tau}^{tv,a}=1)\times P(l_u^{tv,a}=1)\\ &= \alpha_{u,\tau}^{tv,a}\times\beta_u^{tv,a}\\ &= r_a\times\theta^{tv}e^{-\theta^{tv}(\tau-\tau_a)}\times\frac{1}{\sigma_{t_a}\sqrt{2\pi}}e^{-\frac{(x-\mu_{t_a})^2}{2\sigma_{t_a}^2}}\end{aligned}\tag{3}$$

B. SN Appearances

The other type of exposures in the ASIE model is Social Network(SN) appearance. Similar to a TV appearance, the attitude of user u towards a SN appearance b from a friend in the network can be divided into three stages.

1) Cognitive: Like the case of TV appearances, $h_{u,\tau}^{sn,b}$ indicates at time τ, whether u remembers SN appearance b. $h_{u,\tau}^{sn,b}$ obeys Bernoulli distribution with the mean $\alpha_{u,\tau}^{sn,b}$.

$$h_{u,\tau}^{sn,b}=\begin{cases}1 & \text{if } u \text{ remembers } b \text{ at } \tau\\ 0 & \text{otherwise}\end{cases}$$
$$h_{u,\tau}^{sn,b}\sim Bernoulli(\alpha_{u,\tau}^{sn,b})$$

The value of $\alpha_{u,\tau}^{sn,b}$ also relies on two factors. The common factor with TV appearance is the time lapse $\tau-\tau_b$. A tweet posted long time ago has a higher probability to be forgotten and it even cannot be shown on the time line page of u. However there is another factor can balance the time decay. If u and b's author w_b are close friends, u will pay more attention on w_b's posts and has a deeper impression on b. Above all, the value of $\alpha_{u,\tau}^{sn,b}$ is positive correlated to the social link strength $\omega(w_b,u)$ and negative correlated to the time lapse $\tau-\tau_b$. Similarly, it is calculated as

$$\alpha_{u,\tau}^{sn,b}=\omega(w_b,u)\times\theta^{sn}e^{-\theta^{sn}(\tau-\tau_b)}\tag{4}$$

where the exponential function parameter θ^{sn} is learned from data.

2) Affective: This stage considers user u's emotion on b. It is modeled as a coin flip trail $l_u^{sn,b}$, and the success probability is $\beta_u^{sn,b}$.

$$l_u^{sn,b}=\begin{cases}1 & \text{if } u \text{ is impressed by } b.\\ 0 & \text{otherwise}\end{cases}$$
$$l_u^{sn,b}\sim Bernoulli(\beta_u^{sn,b})$$

Whether u is impressed lies on if u is interested in the topic of b, including its hashtag, picture and video. Therefore $\beta_u^{sn,b}$ is generated from a normal distribution, and the mean and variance are determined by its topic t_b.

$$\beta_u^{sn,b}=\frac{1}{\sigma_{t_b}\sqrt{2\pi}}e^{-\frac{(x-\mu_{t_b})^2}{2\sigma_{t_b}^2}}\tag{5}$$

3) Conative: At this stage, influenced by b, user u may repost a tweet or post his own opinion in OSN. $g_{u,\tau}^{sn,b}$ represents whether u is influenced by b. It is drawn from a Bernoulli distribution with mean $\gamma_u^{sn,b}$.

$$g_{u,\tau}^{sn,b}=\begin{cases}1 & \text{if } u \text{ is influenced by } b \text{ at } \tau.\\ 0 & \text{otherwise}\end{cases}$$
$$g_u^{sn,b}\sim Bernoulli(\gamma_u^{sn,b})$$

We define the influence probability which is same with TV appearance, as u is influenced by b when u remembers b and is impressed by b.

$$\begin{aligned}&P(u \text{ is influenced by } b \text{ at } \tau.)\\ &\quad= P(u \text{ remembers } b \text{ at } \tau.)\times P(u \text{ is impressed by } b.)\end{aligned}$$

From the definition, we calculate the $\gamma_{u,\tau}^{sn,b}$:

$$\begin{aligned}\gamma_{u,\tau}^{sn,b} &= P(g_{u,\tau}^{sn,b}=1)\\ &= P(h_{u,\tau}^{sn,b}=1)\times P(l_u^{sn,b}=1)\\ &= \alpha_{u,\tau}^{sn,b}\times\beta_u^{sn,b}\\ &= \omega(w_b,u)\times\theta^{sn}e^{-\theta^{sn}(\tau-\tau_b)}\times\frac{1}{\sigma_{t_b}\sqrt{2\pi}}e^{-\frac{(x-\mu_{t_b})^2}{2\sigma_{t_b}^2}}\end{aligned}\tag{6}$$

C. Appearance Aggregation

In the real situation, to build brand familiarity, TV ads are usually broadcast repeatedly and users in social networks will receive ads messages from their different friends. Advertising research [36] shows potential consumers must be exposed several times before they start to form an opinion about a product or service. Therefore in the ASIE model, impressive appearances are aggregated to activate users taking social actions.

At time τ, the TV stream of u is $S_{u,\tau}^{tv}$, with the size $m_{u,\tau}^{tv}$ which includes all TV appearances displayed before τ. For each $a\in S_{u,\tau}^{tv}$, $P(g_{u,\tau}^{tv,a}=1)$ is the probability of u can be influenced by a, calculated according to (3). His SN stream $S_{u,\tau}^{sn}$, with the size $n_{u,\tau}^{tv}$, contains all SN appearances posted by u's friends and before τ and each $b\in S_{u,\tau}^{sn}$ has an influence probability $P(g_{u,\tau}^{sn,b}=1)$ calculated by (6).

$$S_{u,\tau}^{tv}=\{a|\tau_a<\tau\},\quad S_{u,\tau}^{sn}=\{b|\tau_b<\tau\wedge\ (w_b,u)\in E\}$$

When τ is fixed, $S_{u,\tau}^{tv}$, $S_{u,\tau}^{sn}$ and the influence probability of each appearance are determined. The aggregation process determines how many appearances can influence u [28]. The event that u is influenced by k appearances out of total $m+n$ obeys *Poisson binomial distribution*. The success probability can be calculated as

$$F_{u,\tau}^{m+n}(k)=\sum_{B\in F_k}\prod_{i\in B}p_i\prod_{j\in B^c}(1-p_j)\tag{7}$$

where p_i is influence probability of appearance i. F_k is the set of all subsets of k integers that can be selected from $\{1,2,3,\ldots,m+n\}$. B^c is the complement of B, i.e. $B^c=\{1,2,3,\ldots,m+n\}\setminus B$.

We divide set B into TV appearance set B^{tv} and SN appearance set B^{tv}, i.e. $B=B^{tv}+B^{sn}$. Similarly, $B^c=$

$B^{c,tv} + B^{c,sn}$. Therefore, (7) can represented as

$$F_{u,\tau}^{m+n}(k) = \sum_{B\in F_k} \prod_{i\in B^{tv}} p_i^{tv} \prod_{j\in B^{sn}} p_j^{sn} \prod_{i\in B^{c,tv}} (1-p_i^{tv}) \times \prod_{j\in B^{c,sn}} (1-p_j^{sn}) \quad (8)$$

Thus let $P_u(\tau)$ is the probability that u is activated at τ. Based on (8),

$$P_u(\tau|\theta^{tv},\theta^{sn}) = \sum_{k=1}^{m+n} F_{u,\tau}^{m+n}(k) = \sum_{k=1}^{m+n} \sum_{B\in F_m} \prod_{i\in B^{tv}} p_i^{tv} \prod_{j\in B^{sn}} p_j^{sn} \prod_{i\in B^{c,tv}} (1-p_i^{tv}) \prod_{j\in B^{c,sn}} (1-p_j^{sn}) \quad (9)$$

D. Parameters Inference

With the TV stream, SN stream and social network structure, parameters θ^{tv} and θ^{sn} in the ASIE model can be inferred. Review the information we are given: for a TV appearance a, we know its displayed time τ_a, the topic t_a and the TV rating r_a while for a SN appearance b, we know its topic t_b, the author w_b and the posted time τ_b, also regarded as w_b's activated time. The objective is to learn the value of θ^{tv} and θ^{sn}.

Let τ_u represent the activated time of user u in the real situation, so the ground truth of activation probability is $P_u(\tau_u) = 1$. Therefore the inferring strategy is maximizes the likelihood $P_u(\tau_u)$ of all activated users, which is calculated based on (9) in ASIE model. The log-likelihood function is :

$$ln\mathcal{L}(\theta^{tv},\theta^{sn};P_{u_1}(\tau_{u_1}),P_{u_2}(\tau_{u_2}),\cdots,P_{u_N}(\tau_{u_N})) = \sum_{i=1}^{N} ln\, P_{u_i}(\tau_{u_i}|\theta^{tv},\theta^{sn})$$

Therefore our objective function is

$$\theta^{tv},\theta^{sn} = \text{argmax}\ ln\mathcal{L}(\cdot) = \text{argmax} \sum_{u=1}^{N} ln \sum_{m=1}^{n_{u,\tau}} \sum_{B\in F_m} \Big(\prod_{a\in B^{tv}} p_a^{tv} \prod_{b\in B^{sn}} p_b^{sn} \times \prod_{i\in B^{c,tv}} (1-p_i^{tv}) \prod_{j\in B^{c,sn}} (1-p_j^{sn}) \Big)$$

$$p_i^{tv} = \gamma_{u,\tau}^{tv,i} = r_a \times \theta^{tv} e^{-\theta^{tv}(\tau-\tau_a)} \times \frac{1}{\sigma_{t_a}\sqrt{2\pi}} e^{-\frac{(x-\mu_{t_a})^2}{2\sigma_{t_a}^2}}$$

$$p_j^{sn} = \gamma_{u,\tau}^{sn,j} = \omega(w_b,u) \times \theta^{sn} e^{-\theta^{sn}(\tau-\tau_b)} \times \frac{1}{\sigma_{t_b}\sqrt{2\pi}} e^{-\frac{(x-\mu_{t_b})^2}{2\sigma_{t_b}^2}}$$

$$s.t.\ \theta^{tv},\theta^{sn} \in [0,1], \quad (10)$$

To obtain the MLE of an 2-dimensional vector parameter (θ^{tv}, θ^{sn}), we must solve the following likelihood equation:

$$(ln\mathcal{L})' = 0.$$

It can be proceeded the following equations iteratively until the results converge.

$$\frac{\partial ln\mathcal{L}}{\partial\theta^{tv}} = 0, \quad \frac{\partial ln\mathcal{L}}{\partial\theta^{sn}} = 0$$

When extending the above equations according to (10), we find these are transcendental functions and not solvable in closed form. The approximate solutions will be obtained in the experiment section using available optimization toolkit.

V. EXPERIMENT

To test the effectiveness of ASIE in inferring the social influence of TV ads, extensive experiments have been done on the two real-world advertising campaign datasets introduced in Section 2. In this section, we will first talk about the experiment settings, and then provide the experiments with detailed analysis.

A. Experiment Settings

1) Comparison Methods: There are barely other algorithms can be compared with since we are the first to study the SITE problem. We compare our method against the following baselines:

- ASIE: ASIE is the method proposed in this paper, which predict the probability of users posting tweets influenced by both TV broadcasting and friends in the online social network.
- ASIE-*TV*: A variant of framework ASIE, ASIE-TV, predicts users' behaviors with the TV appearances only.
- ASIE-*SN*: Similar to ASIE-TV, ASIE-SN method use only SN appearances to estimate the activated accounts in social networks.
- *K-nearest neighbors(KNN-SN)*: Classical learning method KNN classifies whether a twitter user will be activated by a majority vote of his k nearest neighbors. Obviously KNN just needs information of SN appearances.
- *Regression(Reg-TV)*: Polynomial regression is utilized to predict users tweeting trends according to the ads aired on TV.

2) Evaluation Measures: To evaluate the performance of all comparison methods, we propose to use the following evaluation metrics:

- *mean absolute error* (MAE): calculate the average of the absolute errors between the prediction results and the ground-truth.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\hat{P}_u(\tau) - P_u(\tau)|\ ,$$

where $P_u(\tau)$ is the probability of u being activated at time τ, $\hat{P}_u(\tau)$ is the predicted value and n is the number of tweets in the testing set. The value of MAE is in the range of $[0, 1]$, and smaller number denotes better performance.

- *mean square error* (MSE): calculate the average squares deviation of the prediction results with regard to the

TABLE III
PREDICTION PERFORMANCE COMPARISON OF DIFFERENT METHODS FOR "CDC TIPS"

Measure	Metrics / methods	Percentage of Training Data									
		10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
MAE	ASIE	**0.3127**	**0.3086**	**0.3087**	**0.3060**	**0.3052**	**0.3048**	**0.3055**	**0.3054**	**0.3050**	**0.3047**
	ASIE-TV	0.3424	0.3381	0.3384	0.3354	0.3347	0.3342	0.3350	0.3348	0.3344	0.3341
	ASIE-SN	0.7550	0.7548	0.7559	0.79540	0.7528	0.7536	0.7522	0.7573	0.7544	0.7539
	KNN-SN	0.8083	0.8010	0.8111	0.8097	0.8092	0.8092	0.8079	0.8076	0.8078	0.8090
	Reg-TV	0.4595	0.4436	0.4240	0.4475	0.4329	0.4197	0.4093	0.4639	0.4670	0.4755
MSE	ASIE	**0.1788**	**0.1735**	**0.1737**	**0.1700**	**0.1689**	**0.1683**	**0.1693**	**0.1691**	**0.1685**	**0.1681**
	ASIE-TV	0.2585	0.2430	0.2433	0.2392	0.2381	0.2374	0.2385	0.2383	0.2376	0.2372
	ASIE-SN	0.5792	0.5788	0.5798	0.5769	0.5751	0.5758	0.5741	0.5823	0.5773	0.5771
	KNN-SN	0.6392	0.6221	0.6341	0.6322	0.6316	0.6316	0.6301	0.6296	0.6299	0.6314
	Reg-TV	0.3384	0.3135	0.2900	0.3189	0.3023	0.2885	0.2771	0.3229	0.3321	0.3424
MAD	ASIE	**0.3117**	**0.3090**	**0.3090**	**0.3074**	**0.3067**	**0.3063**	**0.3069**	**0.3068**	**0.3065**	**0.3061**
	ASIE-TV	0.3419	0.3386	0.3388	0.3357	0.3349	0.3343	0.3352	0.3350	0.3345	0.3340
	ASIE-SN	0.7509	0.7451	0.7231	0.7201	0.7191	0.7188	0.7167	0.7146	0.7139	0.7125
	KNN-SN	0.8490	0.8481	0.8453	0.8426	0.8359	0.8317	0.8268	0.8216	0.8203	0.8200
	Reg-TV	0.5375	0.4907	0.4898	0.5190	0.4807	0.4446	0.4081	0.5774	0.5832	0.4917

ground truth. The value also falls into the interval [0, 1], and smaller value denotes better performance:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{P}_u(\tau) - P_u(\tau))^2 \ .$$

- *median absolute deviation* (MAD): reports the median of the absolute deviations from predictions to the truth data. In our experiment, the prediction result is an activation probability in [0, 1], so the median is also in [0, 1].

3) Setup: The ASIE model learns parameters from training data and predict individual's activation probability. Therefore in the experiment, we first divide each campaign into two phases. Data of the initial phase (Mar. 1 - May 31 for "CDC Tips" and Aug. 1 - Oct. 9 for "Legacy Truth") is used for training and the rest of data (Jun. 1 - Jun. 23 for "CDC Tips" and Oct. 10 - Oct. 31 for "Legacy Truth") is for testing. We then labeled users who posted tweets as positive examples, and their activation probabilities are 1. While their social friends who could see their posts but did not take actions are negative examples, and the probabilities are 0.

Both TV stream $\mathcal{S}^{tv}$ and SN stream $\mathcal{S}^{sn}$ used in the experiments are sorted according to broadcasting time and posted time. We also adjusted all local time to Eastern Standard Time (EST) to make sure the time sequence is correct. The social link strength between users u and v is calculated as Jaccard similarity coefficient.

$$\omega(u, v) = \mathcal{J}(u, v) = \frac{F(u) \cap F(v)}{F(u) \cup F(v)}$$

where $F(u)$ is user u's follower set.

B. Experiment Result

We first study the performance of the proposed ASIE method with different ratio of training data. In each round we choose top 10%, 20%, ..., 100% of training data to learn the parameters, and then use them to predict users activation probability in the testing data. The results obtained by different comparison methods evaluated by the MAE, MSE and MAD metrics are available in Table III and Table IV , where Table III is about the "CDC Tips" dataset, and Table IV is obtained from the "Legacy Truth" dataset. In the results, the parameter *time window length* is fixed as ∞.

1) "CDC Tips": To ASIE itself, the results in Table III show the performance of ASIE improves slightly as more training data is available. For instance, the MAE score obtained by ASIE drops from 0.3127 at 10% training ratio to 0.3047 with full training data; Meanwhile, the MSE score decrease 5% from 0.1788 to 0.1681; And the median of error falls about 0.01 between 10% and 100%. The decrease is not huge because that we divide the data according to time, while TV ratings in different time periods have no regular pattern to follow. Thus when adding more training data, the parameters learned from ASIE have no notable advantages on predicting probabilities.

Comparing with other methods, Table III shows that the ASIE model can consistently outperform other baselines evaluated by all measure metrics. Generally, for the MAE, MSE and MAD metrics, the error scores obtained by ASIE is also the lowest among all the comparison methods consistently for various training ratios. For instance, when the training ratio is 100%, the MSE obtained by ASIE is 0.1681, which is 29.1% smaller than MSE of ASIE-TV, and 70.9% smaller than that obtained by ASIE-SN. The results demonstrate our claim that utilizing information from both TV and social network can depict the advertising campaign's social influence better.

The gaps between ASIE serial methods to other methods are much bigger. With full training data, the MSE score of ASIE is 50.9% smaller than Reg-TV, while ASIE-TV is 30.7% smaller than Reg-TV. When comparing with KNN-SN, the MSE error score of ASIE drops 73.4% and ASIE- SN decreases 8.6%. It means even only using one kind of information, the ASIE model still outperforms other methods.

TABLE IV
PREDICTION PERFORMANCE COMPARISON OF DIFFERENT METHODS FOR "LEGACY TRUTH"

Measure	Metrics	Percentage of Training Data									
	methods	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
MAE	ASIE	**0.2535**	**0.2496**	**0.2476**	**0.2463**	**0.2452**	**0.2442**	**0.2461**	**0.2463**	**0.2456**	**0.2450**
	ASIE-TV	0.3699	0.3716	0.3733	0.3743	0.3751	0.3759	0.3752	0.3740	0.3737	0.3732
	ASIE-SN	0.7308	0.7286	0.7243	0.7212	0.7197	0.7197	0.7219	0.7252	0.7261	0.7264
	KNN-SN	0.8244	0.8267	0.8298	0.8332	0.8350	0.8365	0.8359	0.8376	0.8379	0.8374
	Reg-TV	0.4049	0.9048	0.5350	0.7658	0.6455	0.5522	0.7828	0.6859	0.5416	0.5600
MSE	ASIE	**0.1237**	**0.1226**	**0.1222**	**0.1220**	**0.1219**	**0.1220**	**0.1220**	**0.1220**	**0.1219**	**0.1218**
	ASIE-TV	0.2439	0.2477	0.2516	0.2538	0.2556	0.2573	0.2559	0.2532	0.2526	0.2515
	ASIE-SN	0.5657	0.5628	0.5576	0.5545	0.5531	0.5531	0.5552	0.5586	0.5597	0.5600
	KNN-SN	0.7082	0.7116	0.7167	0.7219	0.7250	0.7279	0.7274	0.7300	0.7308	0.7299
	Reg-TV	0.3950	0.8949	0.5251	0.7559	0.6356	0.5423	0.7729	0.6760	0.5317	0.4501
MAD	ASIE	**0.1962**	**0.1891**	**0.1851**	**0.1823**	**0.1795**	**0.1762**	**0.1818**	**0.1826**	**0.1807**	**0.1790**
	ASIE-TV	0.2244	0.2274	0.2304	0.2318	0.2331	0.2348	0.2334	0.2315	0.2312	0.2306
	ASIE-SN	0.8013	0.7984	0.7965	0.7861	0.7850	0.7849	0.7859	0.7883	0.7886	0.7887
	KNN-SN	0.8705	0.8723	0.8749	0.8766	0.8772	0.8800	0.8792	0.8824	0.8831	0.8834
	Reg-TV	0.4912	0.9900	0.6045	0.8167	0.6791	0.6025	0.8316	0.7271	0.6032	0.5889

We also discover that methods only using information of TV ads (ASIE-TV and Reg-TV) achieve better results than those only utilizing SN information (ASIE-SN and KNN-SN). That is mainly because tweets related to "CDC Tips" campaign is much more influenced by TV ads, which is in agreement with Fig. 2.

2) "Legacy Truth": To ASIE itself, Table IV shows with the rise of training data size, its prediction accuracy inclines slightly. For instance, comparing 10% with 100% of training data, MAE declines 3%, from 0.2535 to 0.2450. Other metrics have the similar decrease. The reason of stable accuracy is the same with "CDC Tips", which both TV ratings and tweets amounts change quickly and irregularly. It is worthy to note that in "Legacy Truth", its error values rise a little during 50% to 60% training data. That is due to the high peak shown in Fig. 4 which represents nearly half of the tweets, and the relation of TV and SN in this time period differs from that in the test data greatly. Therefore the parameters learned from the peak cause more errors of testing data. While with more normal training data input, the values of errors drop down.

Comparing with other algorithms, ASIE model still has a better prediction performance than other algorithms evaluated by all measure metrics. Generally, for the MAE, MSE and MAD metrics, the error scores obtained by ASIE is still the lowest among all baselines consistently with changing training ratios. For instance, the MSE obtained by ASIE is 0.1218, when the training instance ratio is 100%. It is 51.6% smaller than MSE of ASIE-TV, and 78.3% smaller than that of ASIE-SN. The results prove the prediction accuracy of user tweeting trend can be greatly improved by exploiting both TV and social information.

While in "Legacy Truth", the advantage of TV methods (ASIE-TV and Reg-TV) over SN methods (ASIE-SN and KNN-SN) is less obvious than that in "CDC Tips". For instance, with whole training data, the MAE score of ASIE-SN is 1.95 times of ASIE-TV in "Legacy Truth", while in "CDC Tips", this value is 2.25. Moreover, the MAE score of KNN-SN is 1.50 times of that of Reg-TV in "Legacy Truth", while this value in "CDC Tips" is 1.70. It implies using SN information in "Legacy Truth" can get a better prediction than in "CDC Tips". This is also a proof that 'Legacy Truth" campaign conducted a better social marketing.

C. Parameter Analysis

In the above experiment, we assume each appearance has a possibility chance to influence a user regardless of when he was exposed, and the time window parameter is fixed as ∞. However in the real situation, the influence of each appearance will degrade as time passes, and only the most recent appearances shown during the time window will have influence on users.

To study the effect of the time window length, we compare the performance of all methods achieved when this parameter is set at 8, 16, 24, 32, 40, 48, 72 hours respectively. The results obtained on the "CDC Tips" and the "Legacy Truth" datasets are shown in Fig. 6 and Fig. 7 respectively, where three subfigurues correspond to the evaluation metrics: MAE, MSE and MAD.

In the "CDC Tips", with the time window length increases, the error scores obtained by ASIE first increase then drop for the MAE, MSE and MAD metrics. Fig. 3 shows most users cannot get enough exposures until time window length is 12 hours, which causes the high error scores during 8 hours to 16 hours. With the time window length extending, the performance of ASIE improves. Comparing with other algorithm, ASIE outperforms them consistently with changing time window length. For example when the time window length is 72 hours, the MSE of ASIE is 0.1922, which is 9.8% lower than that of ASIE-TV, 68.1% lower than that of ASIE-SN, 73.4% lower than that of KNN-SN and 50.9% lower than that of Reg-TV.

In the "Legacy Truth", error score of ASIE rises slightly with the time window length getting larger, which is caused by the high peak of tweets shown in Fig. 4. With a wider time

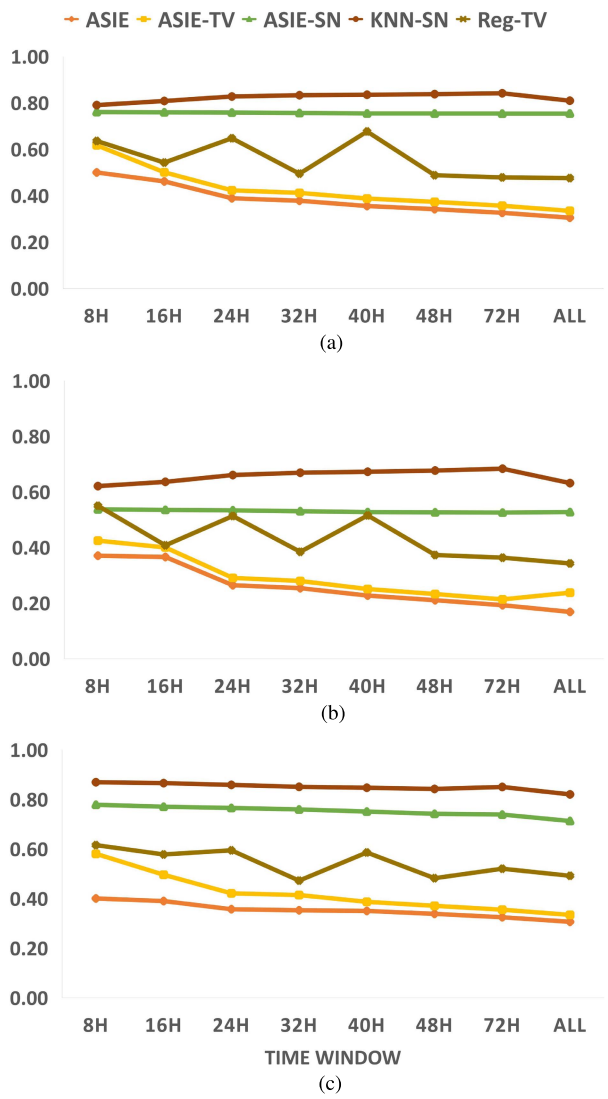

Fig. 6. Performance comparison with different time windows of "CDC Tips". (a) MAE. (b) MSE. (c) MAD.

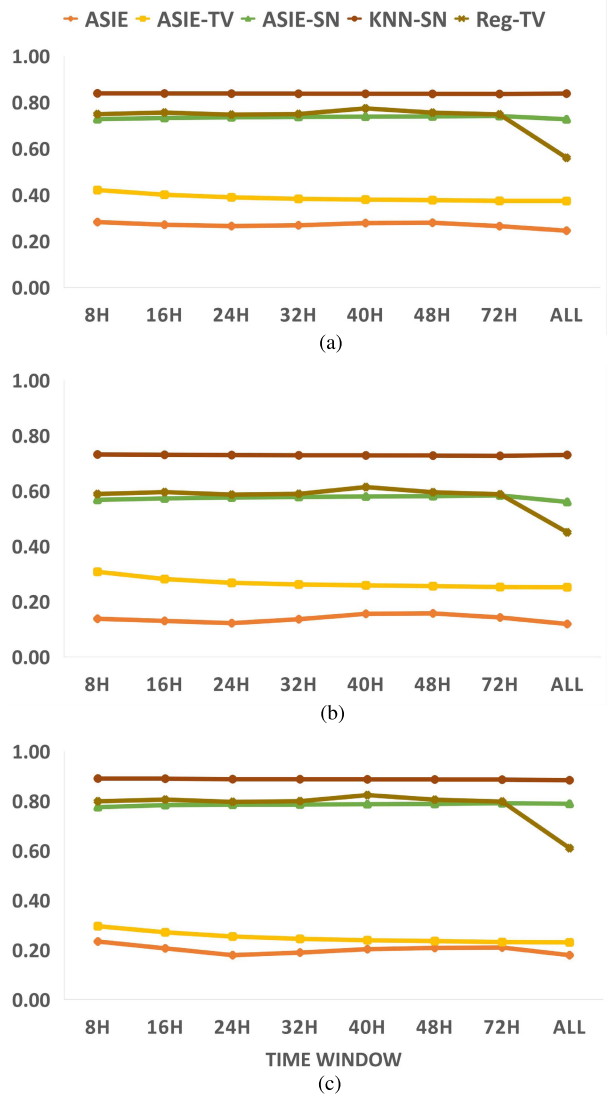

Fig. 7. Performance comparison with different time windows of "Legacy Truth". (a) MAE. (b) MSE. (c) MAD.

window, more and more users are influenced by this extreme condition, which decreases the accuracy of predictions. While comparing with other algorithm, ASIE has the best performance among all methods for various time window lengths.

VI. Related Works

The research on anti-tobacco mass media campaigns attracts scientists in multiple areas, such as public health [13], marketing [30] and communication [18]. Some works [17] [21] also consider the effect of social media in campaigns. While as far as we know, we are the first one in computer science to study the social influence of anti-tobacco campaigns.

In social network analysis area, especially the information diffusion in social networks, has been intensively studied recently [6], [7], [31]. Most of the works regard users in social network are either active or inactive, and information is propagated from active users to inactive ones along the link with a diffusion probability. A plentiful models were constructed to describe this process [12], [19], [23], [27], and among them Independent Cascade(IC) model and its variant [22], [24] were widely used.

Though some research doubt whether diffusion is the only reason activating users [2], [3], [33], most existing works neglect external influence and only focus on internal propagation [20], [25], [37]. There are only two papers paying attention on the effect caused by external trend. [26] proposed an LADP model that improves the learning by extracting social events from data streams. The main step of their method is learning the diffusion probabilities with edges in the networks, and the learned diffusion probabilities can be used in the IC model. While in our model, we are interested how users are activated by public and social media ads, and we do not care about the diffusion probability between each two users.

In the other paper [28], authors identified the role of external out-of-network influence by URL mentioned on Twitter. They presented a model in which information can reach a node via the links of the social network or through the influence of external sources. Our application differs significantly: (1) Our model is based on real-world application of evaluation

advertising campaigns, which is quite different from user influenced by news and social events. Thus we modify the classical psychology and marketing model to describe users' behavior. (2) Their approach focuses on inferring the external trends, while our aims to learn the interaction effect of public and social media ads with the given external information.

VII. Conclusions

In this paper, we propose the SITE problem to predict whether an individual will influenced by anti-tobacco mass media campaigns. We design the ASIE model to infer social influence of the anti-tobacco campaign, which integrates the external TV exposures information and the diffusion process inside the social network. Evaluation on the real TV campaign datasets shows the ASIE model outperforms other baseline methods on the predicting users' tweeting behavior.

References

[1] *Nielsen Reverses Decline in U.S. TV Homes*. Variety. Penske Media Corporation, accessed on May 7, 2013. [Online]. Available: https://en.wikipedia.org/wiki/Nielsen_ratings

[2] S. Aral, L. Muchnik, and A. Sundararajan, "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 51, pp. 21544–21549, 2009.

[3] A. Banerjee, A. G. Chandrasekhar, E. Duflo, and M. O. Jackson, "The diffusion of microfinance," *Science*, vol. 341, no. 6144, p. 1236498, 2013.

[4] T. E. Barry and D. J. Howard, "A review and critique of the hierarchy of effects in advertising," *Int. J. Advertising*, vol. 9, no. 2, pp. 121–135, 1990.

[5] P. Cesar and D. Geerts, "Past, present, and future of social TV: A categorization," in *Proc. IEEE Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2011, pp. 347–351.

[6] M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummadi, "Measuring user influence in twitter: The million follower fallacy," in *Proc. ICWSM*, 2010, vol. 10. nos. 10–17, p. 30.

[7] M. Cha, A. Mislove, and K. P. Gummadi, "A measurement-driven analysis of information propagation in the flickr social network," in *Proc. 18th Int. Conf. World Wide Web*, Apr. 2009, pp. 721–730.

[8] A. Chen, M. Seyoum, R. Panaligan, and K. Wasiljew, *The Role of Digital in Tv Research, Fanship and Viewing*. Think with Google. 2014.

[9] Y.-H. Cheng, C.-M. Wu, T. Ku, and G.-D. Chen, "A predicting model of TV audience rating based on the Facebook," in *Proc. Int. Conf. Social Comput. (SocialCom)*, Sep. 2013, pp. 1034–1037.

[10] K. E. Clow, *Integrated Advertising, Promotion, and Marketing Communications*. Upper Saddle River, NJ, USA: Pearson Education India, 2007.

[11] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach," *Biometrics*, vol. 44, no. 3, pp. 837–845, 1988.

[12] L. Dickens, I. Molloy, J. Lobo, P.-C. Cheng, and A. Russo, "Learning stochastic models of information flow," in *Proc. IEEE 28th Int. Conf. Data Eng. (ICDE)*, Apr. 2012, pp. 570–581.

[13] S. Durkin, E. Brennan, and M. Wakefield, "Mass media campaigns to promote smoking cessation among adults: An integrative review," *Tobacco Control*, vol. 21, no. 2, pp. 127–138, 2012.

[14] S. L. Emery, G. Szczypka, E. P. Abril, Y. Kim, and L. Vera, "Are you scared yet?: Evaluating fear appeal messages in tweets about the tips campaign," *J. Commun.*, vol. 64, no. 2, pp. 278–295, 2014.

[15] E. C. Fieller, H. O. Hartley, and E. S. Pearson, "Tests for rank correlation coefficients. I," *Biometrika*, vol. 44, nos. 3–4, pp. 470–481, 1957.

[16] S. T. Fiske and D. T. Gilbert, *Handbook of Social Psychology*, vol. 2. Hoboken, NJ, USA: Wiley, 2010.

[17] B. Freeman, "New media and tobacco control," *Tobacco Control*, vol. 21, no. 2, pp. 139–144, 2012.

[18] J. Grandpre, E. M. Alvaro, M. Burgoon, C. H. Miller, and J. R. Hall, "Adolescent reactance and anti-smoking campaigns: A theoretical approach," *Health Commun.*, vol. 15, no. 3, pp. 349–366, 2003.

[19] A. Guille and H. Hacid, "A predictive model for the temporal dynamics of information diffusion in online social networks," in *Proc. 21st Int. Conf. Companion World Wide Web*, Apr. 2012, pp. 1145–1152.

[20] A. Guille, H. Hacid, C. Favre, and D. A. Zighed, "Information diffusion in online social networks: A survey," *ACM SIGMOD Rec.*, vol. 42, no. 2, pp. 17–28, 2013.

[21] J. Huang, R. Kornfield, G. Szczypka, and S. L. Emery, "A cross-sectional examination of marketing of electronic cigarettes on twitter," *Tobacco Control*, vol. 23, no. 3, pp. iii26–iii30, 2014.

[22] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2003, pp. 137–146.

[23] A. Khelil, C. Becker, J. Tian, and K. Rothermel, "An epidemic model for information diffusion in MANETs," in *Proc. 5th ACM Int. Workshop Modeling Anal. Simulation Wireless Mobile Syst.*, 2002, pp. 54–60.

[24] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance, "Cost-effective outbreak detection in networks," in *Proc. 13th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2007, pp. 420–429.

[25] S. Lin, Q. Hu, F. Wang, and P. S. Yu, "Steering information diffusion dynamically against user attention limitation," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Dec. 2014, pp. 330–339.

[26] S. Lin, F. Wang, Q. Hu, and P. S. Yu, "Extracting social events for learning better information diffusion models," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2013, pp. 365–373.

[27] Y. Matsubara, Y. Sakurai, B. A. Prakash, L. Li, and C. Faloutsos, "Rise and fall patterns of information diffusion: Model and implications," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2012, pp. 6–14.

[28] S. A. Myers, C. Zhu, and J. Leskovec, "Information diffusion and external influence in networks," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2012, pp. 33–41.

[29] J. Nagy and A. Midha, "The value of earned audiences: How social interactions amplify TV impact," *J. Advertising Res.*, vol. 54, no. 4, pp. 448–453, 2014.

[30] K. Peattie and S. Peattie, "Social marketing: A pathway to consumption reduction?" *J. Bus. Res.*, vol. 62, no. 2, pp. 260–268, 2009.

[31] E. M. Rogers, *Diffusion of Innovations*. New York, NY, USA: Simon and Schuster, 2010.

[32] S. Schirra, H. Sun, and F. Bentley, "Together alone: Motivations for live-tweeting a television series," in *Proc. 32nd Annu. ACM Conf. Human Factors Comput. Syst.*, 2014, pp. 2441–2450.

[33] J. Tang, S. Wu, and J. Sun, "Confluence: Conformity influence in large social networks," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2013, pp. 347–355.

[34] M. A. Wakefield *et al.*, "Impact of tobacco control policies and mass media campaigns on monthly adult smoking prevalence," *Amer. J. Public Health*, vol. 98, no. 8, pp. 1443–1450, 2008.

[35] K. Weller, A. Bruns, J. Burgess, M. Mahrt, and C. Puschmann, *Twitter and Society*. New York, NY, USA: Peter Lang, 2013.

[36] S. Moriarty, N. D. Mitchell, W. D. Wells, R. Crawford, L. Brennan, and R. Spence-Stone, *Advertising: Principles and Practice*. Pearson Australia, 2014.

[37] Q. Zhan, J. Zhang, S. Wang, S. Y. Philip, and J. Xie, "Influence maximization across partially aligned heterogenous social networks," in *Proc. Adv. Knowl. Discovery Data Mining*, 2015, pp. 58–69.