Temporal Causality Analysis of Sentiment Change in a Cancer Survivor Network

Ngot Bui, John Yen, *Fellow, IEEE*, and Vasant Honavar, *Senior Member, IEEE*

***Abstract*— Online health communities (OHCs) constitute a useful source of information and social support for patients. American Cancer Society's Cancer Survivor Network (CSN), a 173 000-member community, is the largest online network for cancer patients, survivors, and caregivers. A discussion thread in CSN is often initiated by a cancer survivor seeking support from other members of CSN. Discussion threads are multiparty conversations that often provide a source of social support, e.g., by bringing about a change of sentiment from negative to positive on the part of the thread originator. While previous studies regarding cancer survivors have shown that the members of an OHC derive benefits from their participation in such communities, causal accounts of the factors that contribute to the observed benefits have been lacking. We introduce a novel framework to examine the temporal causality of sentiment dynamics in the CSN. We construct a probabilistic computation tree logic representation and a corresponding probabilistic Kripke structure to represent and reason about the changes in sentiments of posts in a thread over time. We use a sentiment classifier trained using machine learning on a set of posts manually tagged with sentiment labels to classify posts as expressing either positive or negative sentiment. We analyze the probabilistic Kripke structure to identify the prima facie causes of sentiment change on the part of the thread originators in the CSN forum and their significance. We find that the sentiment of replies appears to causally influence the sentiment of the thread originator. Our experiments also show that the conclusions are robust with respect to the choice of the: 1) classification threshold of the sentiment classifier and 2) choice of the specific sentiment classifier used. We also extend the basic framework for temporal causality analysis to incorporate the uncertainty in the states of the probabilistic Kripke structure resulting from the use of an imperfect state transducer (in our case, the sentiment classifier). Our analysis of temporal causality of CSN sentiment dynamics offers new insights that the designers, managers, and moderators of an online community, such as CSN, can utilize to facilitate and enhance the interactions so as to better meet the social support needs of the CSN participants. The proposed methodology for the analysis of temporal causality has broad applicability in a variety of settings where the dynamics of the underlying system can be modeled in terms of state variables that change in response to internal or external inputs.**

***Index Terms*— Online health community (OHC), sentiment classification, sentiment dynamics, temporal causality.**

I. INTRODUCTION

WORLD health organization [1] estimated that 14.1 million new cancer cases and 8.2 million cancer-related deaths occurred worldwide in 2012. In 2014, the number of deaths due to cancer in the U.S. was estimated to be in excess of 0.58 million, and the number of new cancer cases diagnosed was estimated to be 1.66 million [2]. According to the National Cancer Institute, as of January 1, 2012, there were approximately 13.7 million cancer survivors in the U.S. [2]. While some of the cancer survivors were cancer free, others continued to exhibit cancer symptoms for which they were possibly being treated [2].

According to a Pew Research study, 72% of Internet users in the U.S. utilize the Internet for health-related purposes and 26% have read or watched someone else's experience about health or medical issues during the previous year [3]. Online health communities (OHCs) constitute an important source of information and social support for patients [4] beyond that available through family members, friends, or even health care providers [5]. Previous studies have found that the OHC participants report increased social support [4], [6], reduced levels of depression, stress, and psychological trauma [7], increased optimism about life [6], increased ability to cope with patients' health conditions, and improvements in both the physical and the psychological aspects of lives [4], [8], [9].

American Cancer Society's Cancer Survivor Network (CSN), a 173 000-member community, is the largest online network for cancer patients, survivors, and caregivers. A discussion thread in CSN is often initiated by a cancer survivor seeking support from other members of CSN. Such discussion threads are multiparty conversations that often provide a source of social support, e.g., by bringing about a change of sentiment from negative to positive on the part of the thread originator. Several studies have used sentiment analysis of posts in an OHC to examine the benefits of such interactions to the participants [10], [11] as well as for assisting moderators to intervene in ways that increase the benefits derived by the participants [12]. Sentiment analysis of content in the breast and colorectal cancer forums of the CSN has been used to examine how the change in sentiment of thread originators depends on the topic of the discussion thread [10] and to study the benefits

Manuscript received June 13, 2015; revised January 15, 2016 and May 23, 2016; accepted June 30, 2016. Date of publication August 10, 2016; date of current version November 23, 2016. This work was supported by a collaborative agreement with the American Cancer Society, which made the data of CSN available for this Research, in part by the Edward Frymoyer Endowed Professorship in Information Sciences and Technology (held by V. Honavar), in part by the National Science Foundation (NSF) under Grant CNS 1551843, and in part by the National Center for Advancing Translational Sciences, National Institutes of Health (NIH), through Grant UL1 TR000127. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NSF or the NIH. The authors would like to thank K. Portier, G. Greer of the American Cancer Society, current and former members of the Artificial Intelligence Research Laboratory (led by V. Honavar), Applied Cognitive Science Laboratory (led by D. Reitter), and the Cancer Informatics Group (led by J. Yen) at Pennsylvania State University for the useful discussions.

The authors are with the College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA 16802 USA (e-mail: npb123@ist.psu.edu; jyen@ist.psu.edu; vhonavar@ist.psu.edu).

Digital Object Identifier 10.1109/TCSS.2016.2591880

of social support and information support provided by online interactions [13]. Adequately addressing the social support needs of participants in an OHC, such as CSN, calls for a causal account of aspects of online interaction that enhance the benefits of participation, e.g., an improved sense of wellbeing, as indicated, for example, by a positive change in the sentiment. However, with the exception of our recent work [14], none of the preceding studies have provided a causal account of sentiment change on the part of OHC participants.

Causal Bayesian networks [15]–[17], a family of structural causal models, offer a powerful approach to representing, reasoning about, and eliciting causal relationships from observational and experimental data. Structural causal models offer a synthesis of graphical models that encode scientific assumptions in a qualitative (nonparametric) language on the one hand and the frameworks of counterfactuals [18] and potential outcomes [19] on the other. Dynamic Bayesian networks [20]–[22] extend causal Bayesian networks to allow the modeling of such causal relationships in dynamical systems, where in the states of some variables of interest, e.g., one's current education level and current work habits causally influence the future states of other variables of interest, e.g., one's future income. Granger causality [23] offers a framework for determining whether one time series provides useful information in predicting another [time series A is said to Granger cause B if it can be shown that time-lagged values of A provide statistically significant information about future values of B (beyond that provided by time-lagged values of B)]. However, Granger causality does not, strictly speaking, capture causal relationships; and dynamic Bayesian networks lack the expressivity needed to represent and reason about the temporal properties of the underlying system, e.g., the first positive reply to a post expressing a negative sentiment, with at least 80% probability, results in a change in sentiment within 5 hours. Against this background, we introduce a novel approach that leverages the machinery of temporal causality developed in [24] to uncover the temporal causality of the dynamics of sentiment change (on the part of the thread originators) in OHCs. This approach explicitly captures the temporality of the relationship between cause and effect. In addition to being able to represent properties being true for durations of time, it also allows a direct representation of the time it takes for the cause to show its effect.

Specifically, we introduce a novel framework to examine the temporal causality of sentiment dynamics in OHCs. We demonstrate our methodology using CSN data. We use a probabilistic computation tree logic (PCTL) encoding and a corresponding probabilistic Kripke structure to represent and reason about the changes in sentiments of posts in a thread over time. We use a sentiment classifier trained using machine learning on a set of posts manually tagged with their sentiment labels to classify posts as expressing either positive or negative sentiment. We analyze the probabilistic Kripke structure to identify the prima facie causes of sentiment change on the part of the thread originators in the CSN forum and their significance. Our analysis, perhaps not surprisingly, shows that the sentiment of replies appears to causally influence the sentiment of the thread originator. These conclusions are robust with respect to the choice of the: 1) classification threshold of the sentiment classifier and 2) choice of the specific sentiment classifier used. We extend our basic framework for temporal causality analysis to incorporate the uncertainty in the states of the probabilistic Kripke structure resulting from the use of an imperfect state transducer (in our case, the sentiment classifier). Our methodology can be used to gain new insights that the designers, managers, and moderators of an online community, such as CSN, can utilize to facilitate and enhance the interactions so as to better meet the social support needs of the CSN participants. The methodology for analysis of temporal causality has broad applicability in a variety of settings where the dynamics of the underlying system can be modeled in terms of state variables that change in response to internal or external inputs.

This paper substantially extends our previous work [14] on temporal causality on social support in OHCs along several directions. First, we analyze the temporal causality relationship as a function of the discussion topic. Second, we investigate the robustness of the framework with respect to the choice of the classification threshold of the sentiment classifier and the choice of the specific sentiment classifier used. Third, we offer a modification of the basic framework for temporal causality analysis and, hence, uncertainty in the states of the probabilistic Kripke structure resulting from the use of an imperfect state transducer (e.g., an imperfect sentiment classifier in an analysis of temporal causality of sentiment dynamics).

The rest of this paper is organized as follows. Section II introduces the key notions of temporal causality and the machinery for reasoning about causes of events. Section III describes our approach in temporal causality analyzing in CSN. Section IV presents the results of experiments that demonstrate the utility of the proposed approach. Section V investigates the robustness of the inferred temporal causality of CSN sentiment dynamics with regard to the choice of classification threshold. Section VI shows how to incorporate sentiment classification error rates into the analysis of temporal causality of CSN sentiment dynamics. Section VII examines the robustness of the inferred temporal causality of CSN sentiment dynamics with regard to the choice of the specific sentiment classifier used. Section VIII concludes with a summary and a discussion and an outline of some promising directions for further research.

II. Temporal Causality

We start by reviewing a few key notions [25]. An event B is said to be a prima facie or potential cause of an event A if and only if: 1) B precedes A; 2) the probability of B, $p(B) > 0$; and 3) the conditional probability $p(A|B) > p(A)$. We say that B is a spurious cause of A if and only if B is a prima facie cause of A, and there is an event C that precedes B such that: 1) $p(B, C) > 0$; 2) $p(A|B, C) = p(A|C)$; and 3) $p(A|B, C) \geq p(A|B)$. That is, C occurs before B and accounts for the effect A as well as B does. For example, assume that smoking and yellow finger precede the development of lung cancer and both appear to be the causes to lung cancer. However, yellow finger and lung cancer have the same common cause (i.e., smoking). A prima facie cause that is not a spurious cause is said to be a genuine cause [25]. Suppes [25] offers a method for testing whether a cause is spurious in the restricted setting where there are only two possible causes. Kleinberg [24] argued for a more stringent criterion, for a prima facie cause to be considered a genuine cause and introduced a method for assessing the causal significance of a potential cause of an effect, which can be used to identify a genuine cause of an event from among a set of its potential causes.

Kleinberg's framework [24] uses temporal logic [26] to represent and reason about events that occur in time. Temporal logic is a variant of propositional modal logic that admits the truth value of a formula, constructed from atomic propositions (sentences that are either true or false and encoded by propositional symbols) using logical connectives (i.e., conjunction, disjunction, and negation) to be time-dependent. Hence, temporal logic can be used to represent whether a property is true at some specific time. CTL [27], a branching-time logic, can be used to represent the fact that a property will be true at some time in the future (e.g., at some point in the future, the train will arrive). PCTL [28] extends CTL by specifying deadlines (requiring a property to hold before a specified window of time elapses) and quantifying the transition probability between the states in CTL. Probabilistic Kripke structures [27] can be used to represent and reason in PCTL.

Definition 1 (Probabilistic Kripke Structure): Let AP be a set of atomic propositions, and a probabilistic Kripke structure is a four tuple: $K = \langle S, s^i, L, \mathcal{T} \rangle$, where S is a finite set of states; $s^i \in S$ is an initial state; $L : S \to 2^{AP}$ is a labeling function assigning subsets of AP to states; and $\mathcal{T} : S \times S \to [0, 1]$ is a transition probability function, such that $\forall s \in S : \sum_{s' \in S} \mathcal{T}(s, s') = 1$.

PCTL [28] provides three types of modal operators that incorporate probabilities into their CTL counterparts: A (for all paths) and E (for some future path), temporal operators: G (holds for entire future path) and F (finally holds), and the leads-to operator. The leads-to operator, which is useful in formalizing temporal priority for causality, is defined as $h \leadsto^{\leq t}_{\geq p} g \equiv AG[h \to F^{\leq t}_{\geq p} g]$, which means that whenever h holds, there is a probability of at least p that g will hold via some series of transitions taking less than or equal t time units. Some scenarios require the specification of a lower bound of time for g to hold. In this case, the leads-to operator can be constructed as $h \leadsto^{\geq t_1, \leq t_2}_{\geq p} g$, which denotes that g must hold in between t_1 and t_2 time units with probability p where $0 \leq t_1 \leq t_2 \leq \infty$ and $t_2 \neq \infty$.

Armed with the machinery of PCTL, we can define a prima facie (or a potential) cause as follows [29].

Definition 2 (Prima Facie Cause in PCTL): c is a prima facie cause of e if there is a p such that all three following conditions hold: 1) $F^{\leq \infty}_{>0} c$; 2) $c \leadsto^{\geq 1, \leq \infty}_{\geq p} e$; and 3) $F^{\leq \infty}_{<p} e$, where c and e are PCTL formulas.

For c to be a prima facie cause of an effect e, the state where c is true should be reachable with nonzero probability, and the probability of reaching a state where e is true from a state where c is true should be greater than the probability of reaching a state where e is true from the initial state of the system. This can be interpreted as requiring that c must occur at least once, and that the conditional probability of e given c is greater than the marginal probability of e.

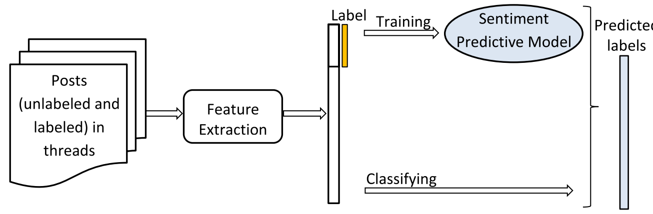


Fig. 1. Classifying sentiments of CSN posts.

We adopt the technique from [28] to calculate the probabilities $F^{\leq \infty} c$, $c \leadsto^{\geq 1, \leq \infty} e$, and $F^{\leq \infty} e$, where $F^{\leq \infty} e$ denotes the path probabilities summed over the set of all paths starting from the initial state of the Kripke structure K and ending in a state where e is true (i.e., the unconditional probability of e); $c \leadsto^{\geq 1, \leq \infty} e$ denotes the path probabilities summed over the set of all paths starting from the state where c is true and ending at a state where e is true (i.e., the conditional probability of e given c).

III. Methodology

A. Sentiment Classification of CSN Posts

Since we cannot detect the sentiment of a CSN member directly, we use the sentiment expressed in a post as a proxy for the sentiment of the CSN member at the time the post was created. As in [30], we categorize the posts as expressing either a positive sentiment or a negative sentiment. To assign sentiment labels to the posts, we use a sentiment classifier that is trained on a subset of posts for which the sentiment labels are assigned manually. The training data for our sentiment classifier were provided to us in [11] wherein the sentiment labels were manually assigned by a rater (a graduate student) and reviewed by experts from the American Cancer Society. We used the labeled data to train our sentiment classifier using machine learning and used the resulting classifier to assign sentiment labels to the unlabeled posts in the larger data set. In our analyses, we used the resulting data set of posts labeled with sentiment labels.

Fig. 1 shows the procedure used to train the post sentiment classifier and to apply it to classify the new posts according to their sentiment. We used a random sample of 298 posts, which was selected from the CSN breast cancer forum and each post manually classified as being of positive or negative sentiment with the result that 204 of them were labeled as positive and 94 were negative (this training data set is obtained from [11]). As in [11], we extracted seven features (see Table I) from a post to train a predictor for assigning posts to the positive or negative category. SentiStrength [30] is used to extract PosStrength and NegStrength, which represent the positive sentiment strength and negative sentiment strength of the post, respectively. We make use of the four lists,[1] a list of positive and negative words [31], a list of positive and negative emotions,[2] a list of popular English female and male names,

[1] http://sites.google.com/site/qiubaojun/psu-sentiment.zip

[2] http://en.wikipedia.org/wiki/List_of_emotions

TABLE I
FEATURES OF A POST

Feature Name	Description
PostLength	The number of words in the post
PosStrength	Positive Sentiment Strength of the post
NegStrength	Negative Sentiment Strength of the post
Neg	NumberOfNeg / PostLength, where NumberOfNeg is equal to number of negative words and symbols, e.g., ":("
PosVsNeg	(NumberOfPos + 1) / (NumberOfNeg + 1), where NumberOfPos is equal to number of positive words and symbols
Name	NumberOfName / PostLength, where NumberOfName is the number of names mentioned in the post
Slang	Number of Internet slang words in the post

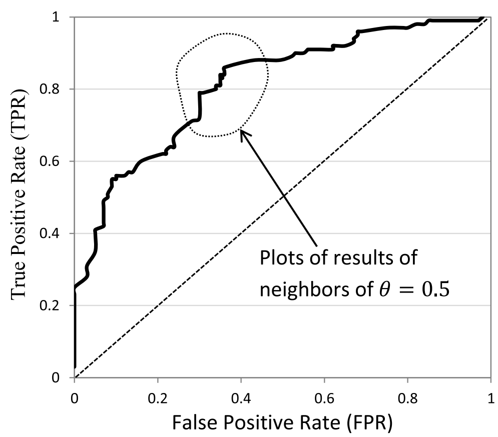


Fig. 2. ROC of Adaboost classifier.

and a list of Internet slang words to calculate the Neg, PosVsNeg, Name, and Slang features. We used Adaboost (regression trees are used as weak learners), which has been shown to generate the best performing sentiment classification model [11] to classify the posts. The Adaboost sentiment classifier outputs a probability that a post expresses a positive sentiment, i.e., $p(\text{positive} \mid \text{post})$. If $p(\text{positive} \mid \text{post}) > \theta$, the post is classified as positive; otherwise, it is classified as negative. Inspection of the Receiver Operating Characteristic (ROC) curve for the classifier (see Fig. 2) shows that thresholds in the neighborhood of $\theta = 0.5$ yield an optimal tradeoff between false positive and true positive predictions. Hence, in our experiments, we used a classification threshold of $\theta = 0.5$. The resulting classifier has an area under the ROC curve $= 0.832$ and a classification accuracy $= 79.2\%$ (as estimated by 10-fold cross-validation).

B. CSN Thread Viewed as a Sequence of Sentiments

The CSN operated by the American Cancer Society is an OHC with over 173 000 registered members, which include cancer patients, their friends and families, and informal caregivers. In this paper, we use the CSN data set that contains all threads initiated between July 2000 and October 2010. The data set contains 48 779 discussion threads and more than 468 000 posts from 27 173 users. The data set is appropriately anonymized to protect the privacy of the CSN members.

Our goal is to uncover the causal effect (if any) of the temporally ordered posts that make up the thread on the final sentiment of the thread originators. More specifically, we are interested in discovering causal relationship between the reply posts and the change of sentiment of those who initiate the thread. Therefore, threads used in this paper need to have at least one reply and at least one self-reply (i.e., a post by the thread originator later on the thread). As a result, threads that do not contain a self-reply or reply are not considered in our analyses. The resulting data set consists of 22 854 threads (Table II shows the distribution of threads over years).

A thread can be represented as a temporally ordered sequence of posts $P_{o1}, P_{r1}, P_{r2}, \ldots, P_{o2}, \ldots, P_{rm}, P_{on}$, where P_{o1} is the initial post from the thread originator; P_{oi} $(i > 1)$ are self-replies; and P_{rj} are replies to the post by individuals other than the thread originator. Since we focus on the communication between two kinds of actors in a thread, the thread originator and the individuals (other than the originator) who respond to the originator's post, we simply compute the average probability of positive sentiment of replies between two consecutive self-replies and use it as the positive sentiment probability of the collection of replies. Formally, the average positive sentiment probability is calculated as

$$\bar{p}_{ri} = \frac{\sum_j p(\text{positive} \mid P_{rj})}{N_{ri}}, \; j : t_{oi} \le t_{rj} \le t_{o(i+1)} \qquad (1)$$

where t_{oi}, $t_{o(i+1)}$, and t_{rj} are time points when posts P_{oi} and $P_{o(i+1)}$ and P_{rj} are created, respectively, and N_{ri} is the number of reply posts from t_{oi} to $t_{o(i+1)}$.

We transform the sequence of post sentiment probabilities in a thread to a sequence of post sentiment states as follows: [Sentiment state of initial post] [Average sentiment state of reply posts] ([Sentiment state of intermediate self-reply] [Average sentiment state of reply posts])* [Sentiment state of final self-reply], where average sentiment state of reply posts is obtained from the average sentiment probability defined in formula (1) using the threshold of sentiment state classifier described in Section III-A (i.e., $\theta = 0.5$). More precisely, each sentiment state can take one of two values: positive or negative. Let b, o, r, s, and f be atomic propositions. Let b denote the beginning of a thread, o denote that the initial post sentiment is positive, r denote that the average sentiment of reply posts elicited by the initial post is positive, s denote that the sentiment of an intermediate self-reply to the initial post is positive, and f denote that the sentiment of the final self-reply is positive. A thread can be represented by a sentiment state sequence $\mathbf{x} = x_0 x_1 \ldots x_n$, where $x_i \in \mathcal{X} = \{o, \neg o, r, \neg r, s, \neg s, f, \neg f\}$, where $\neg$ denotes the negation of the corresponding proposition.

C. Probabilistic Kripke Structure for CSN Sentiment Dynamics

We use a probabilistic Kripke structure [27] to represent and reason about probabilistic transitions between the sentiment states of posts in a CSN. Let $\mathbf{x} = x_0 x_1 \ldots x_n$ be a sequence of post sentiments where $x_i \in \mathcal{X}$ and let $X_i (0 \le i \le n)$ be a random variable corresponding to a sequence element x_i. Markov model (MM), which captures the dependencies

TABLE II
NUMBER OF THREADS VERSUS YEAR

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Number of threads	79	316	458	602	1080	962	1330	1146	1646	7219	8016

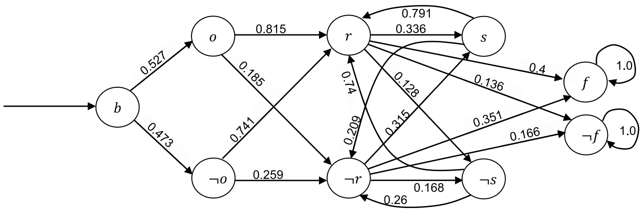


Fig. 3. Probabilistic Kripke structure for CSN sentiment change.

TABLE III
SENTIMENT CHANGE IN CSN

$\#[o]$	$\#[\neg o]$	$\#[\neg o \rightarrow f]$	$\#[o \rightarrow \neg f]$
12147	10707	7512	2948

between the neighboring sequence elements, is used to estimate the transition probabilities between sentiment states of posts that make up a thread. In the kth order MM, the sequence element follows the Markov property: $X_i \perp\!\!\!\perp \{X_0, \ldots, X_{i-k-1}\} \mid \{X_{i-k}, \ldots, X_{i-1}\}$ (i.e., X_i is conditionally independent of $X_0, \ldots, X_{i-k-1}$ given $X_{i-k}, \ldots, X_{i-1}$ for $i = k, \ldots, n$). Formally, the transition probabilities are estimated[3] over a set $\mathcal{D} = \{\mathbf{x}_l\}_{l=1}^{|\mathcal{D}|}$ of sentiment sequences as follows:

$$\hat{p}(X_i = \sigma | w) = \left[\frac{1 + \sum_{l=1}^{|\mathcal{D}|} \#[w\sigma, \mathbf{x}_l]}{|\mathcal{X}| + \sum_{\sigma' \in \mathcal{X}} \sum_{l=1}^{|\mathcal{D}|} \#[w\sigma', \mathbf{x}_l]} \right]_{\sigma \in \mathcal{X}} \tag{2}$$

where $\#[w\sigma, \mathbf{x}_l]$ represents the number of times the symbol σ follows the subsequence w (of length k) in the sequence $\mathbf{x}_l$ and $\hat{p}(X_i = \sigma | w)$ is the estimate of the conditional probability $p(X_i = \sigma | w)$ of the sequence element X_i appearing after the subsequence w. We use the first-order MM to determine the transition probabilities for the CSN probabilistic Kripke structure.

IV. TEMPORAL CAUSALITY OF SENTIMENTS IN CSN

A. Prima Facie Cause

Fig. 3 shows the Probabilistic Kripke structure (K) that is constructed using the method described in Section III-C. The structure shows that from any state of the thread originator, i.e., $\{o, \neg o, s, \neg s\}$, there is a probability of at least 74% that it will transit to the state r. This suggests that CSN participants who respond to thread originators tend to express positive sentiment regardless the sentiment of the thread originators. In other words, members of CSN try to offer positive social support to CSN participants who initiate a thread seeking support from the community. Table III shows that about 70.2% of thread originators with initial negative sentiment end up with positive sentiment at the end of the thread, and about 24.3% of thread originators with initial positive sentiment end up with a negative sentiment at the end of the thread.

Our goal is to uncover the prima facie causes for final sentiment of the thread originators. Based on the definition of prima facie causes and the probabilistic Kripke structure K, we find that the set of prima facie causes of f and $\neg f$ are $\{r, s\}$ and $\{\neg o, \neg r, \neg s\}$, respectively. We adopt the technique from [28] to calculate the probabilities $F^{\leq\infty}c$, $c \rightsquigarrow^{\geq 1, \leq \infty} e$, and $F^{\leq\infty}e$ in case of the probabilistic Kripke structure K. Suppose we want to find out all the potential causes of the positive sentiment f. Let $p(\infty, c)$ be the path probabilities summed over the set of all paths starting from the state where c is true and ending at a state where e is true. According to definition 2, we have $F^{\leq\infty} f = p(\infty, b)$ (b is the initial state of the probabilistic Kripke structure K) and $c \rightsquigarrow^{\geq 1, \leq \infty}_{\geq p} f = p(\infty, c)$. The recursive formula for $p(\infty, b)$ is as follows: $p(\infty, b) = \mathcal{T}(b, o) \times p(\infty, o) + \mathcal{T}(b, \neg o) \times p(\infty, \neg o)$, where $p(\infty, o) = \mathcal{T}(o, r) \times p(\infty, r) + \mathcal{T}(o, \neg r) \times p(\infty, \neg r)$, and so on. Unrolling the recursion yields solutions for $p(\infty, b)$ and $p(\infty, c)$ where $c \in \{o, \neg o, r, \neg r, s, \neg s\}$. For each c, if there exists a probability p that satisfies the three prima facie conditions in definition 2, c is a prima facie cause of f. For example, with the probabilistic Kripke structure in Fig. 3, we find that $p(\infty, b) = F^{\leq\infty} f = 0.73$ and $p(\infty, r) = 0.74$ and $p(\infty, \neg r) = 0.7$, which means that r is prima facie cause of f but $\neg r$ is not a prima facie cause of f.

CSN is comprised of users' data over a period of 11 years. CSN can be divided into several subcommunities (e.g., breast cancer and colorectal cancer). To validate the above prima facie causes, we divided CSN data set into several subsets based on the year and the subcommunity.

Specifically, we group the threads that originated in the same year (from 2000 to 2010) and we group the threads that belong to either Breast Cancer or Colorectal Cancer Community (from 2005 to 2010). Surprisingly, r and $\neg r$ are consistently the prima facie causes of f and $\neg f$, respectively, in both yearly and subcommunity data sets. Tables IV and V show the prima facie causes of f and $\neg f$ in the yearly and subcommunity analysis. The results from the two tables indicate that the positive sentiment of the replies appears to causally influence the positive sentiment of the thread originator at the end of the thread; conversely, the negative sentiment of the replies appears to causally influence the negative sentiment of the thread originator at the end of the thread.

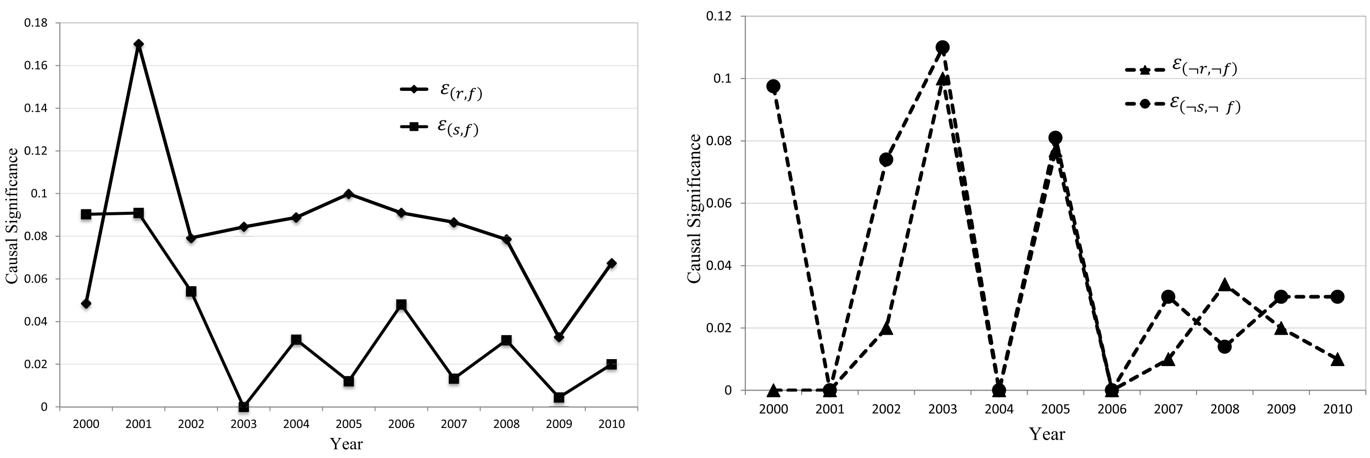
B. Assessing Causal Significance

We proceed to evaluate the significance of the prima facie causes of f and $\neg f$ using the method introduced in [29]. We calculate the causal significance of a prima facie cause c for an effect e as $\varepsilon(c, e) = (\sum_{x \in X \setminus c} \varepsilon_x(c, e) / |X \setminus c|)$, where X is a set of prima facie causes of e and $\varepsilon_x(c, e) = p(e | c \wedge$

[3]Laplace correction is used for smoothing the estimates.

TABLE IV
YEARLY PRIMA FACIE CAUSES OF FINAL THREAD ORIGINATOR'S SENTIMENT

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
f	o	o, r	r, s	r, s	r, s	r, s	r, s	r, s	r	r, s	r, s
$\neg f$	$\neg o$	$\neg o, \neg r, \neg s$	$\neg o, \neg r$	$\neg r, \neg s$	$\neg r, \neg s$	$\neg r, \neg s$	$\neg r, \neg s$	$\neg r, \neg s$	$\neg r$	$\neg r, \neg s$	$\neg r, \neg s$



(a) (b)

Fig. 4. Causal significance by year. (a) Causal significance of r and s with respect to f. (b) Causal significance of $\neg r$ and $\neg s$ with respect to $\neg f$.

TABLE V
PRIMA FACIE CAUSES OF FINAL THREAD ORIGINATOR'S SENTIMENT CHANGE

Community	Breast Cancer	Colorectal Cancer
f	r	r
$\neg f$	$\neg r$	$\neg r, \neg s$

TABLE VI
CAUSAL SIGNIFICANCE

$\varepsilon(r, f)$	$\varepsilon(s, f)$	$\varepsilon(\neg o, \neg f)$	$\varepsilon(\neg r, \neg f)$	$\varepsilon(\neg s, \neg f)$
0.054	0.01	0.05	0.04	0.039

$x) - p(e|\neg c \wedge x)$ denotes the contribution of c to the change in probability of e. Table VI shows causal significance between causes and effects from an aggregate of data from all the years.

From Table VI, we can see that causal significance $\varepsilon(r, f)$ is much higher than the causal significance $\varepsilon(s, f)$ and whereas $\varepsilon(\neg o, \neg f)$, $\varepsilon(\neg r, \neg f)$, and $\varepsilon(\neg s, \neg f)$ are not much different from each other.

In a similar fashion, we also examined the causal significance on data for specific years and subcommunities. Fig. 4 shows the results of this analysis. Fig. 4(a) shows that the causal significance $\varepsilon(r, f)$ is significantly greater (paired t-test, $p < 0.01$) than the causal significance $\varepsilon(s, f)$. However, Fig. 4(b) shows that $\varepsilon(\neg r, \neg f)$ and $\varepsilon(\neg s, \neg f)$ are not significantly different [Fig. 4(b) does not include the significance of $\neg o$, since it is not found to be a cause of $\neg f$ in most of the years (except during the first three years, which account for less than 4% of the total number threads)]. Our analysis of the data from the subcommunities yields a similar finding [i.e., $\varepsilon(r, f)$ is significantly greater than $\varepsilon(s, f)$, and $\varepsilon(\neg r, \neg f)$ and $\varepsilon(\neg s, \neg f)$ are not significantly different from each other].

Based on the results summarized in Table VI and Fig. 4, we can conclude that r causally influences f and $\{\neg r, \neg s\}$ causally affect $\neg f$. In other words, our key finding is that the positive sentiment of a reply causes the negative to positive change in the thread originator's sentiment, at least among 70.2% of the thread originators with an initial negative sentiment. We also see that negative sentiment from a replier causes the thread originator to be left with a negative sentiment. Hence, we conclude that the sentiment of the replies drives the sentiment dynamics of the thread originator.

C. Temporal Causality as a Function of Discussion Topic

In Sections IV-A and IV-B, we analyzed the temporal causality of the CSN data set and several of its subsets (i.e., corresponding to the different CSN subcommunities, and the years). Here, we explore the temporal causality of sentiment dynamics as a function of the discussed topic within a thread. This analysis is motivated in part by an earlier study in [32], which identified two kinds of support received by CSN participants: emotional support and informational support. To investigate the temporal causality of sentiment dynamics as a function of the discussion topic, we perform topic analysis[4] of the CSN threads using LDA [33]. We use the results of topic analysis to identify the popular topics within the CSN and then categorize threads by topics. Because the output of LDA is a soft clustering of threads, a thread can be assigned to multiple topics. We then proceed to perform

[4]Note that one can perform topic assignment at post-level. However, posts in CSN are often quite short. Current methods for topic modeling are not particularly effective on such short documents.

TABLE VII
POPULAR TOPICS IN CSN DISCUSSION THREADS

"Action" (14567)	"Feeling" (11699)	"Time" (9715)	"Family" (8160)	"Emotion" (7777)	"Sickness" (7722)	"Treatment" (7484)	"Finance" (5606)
GET	PRAYER	CHEMO	MOM	HAPPY	CANCER	SCAN	INSURANCE
TAKE	WILL	TREATMENT	HUSBAND	NEWS	SMOKER	CT	WORK
LIKE	GOOD	RADIOACTIVE	DAD	GOOD	STAGE	TEST	PAY
KNOW	HOPE	WEEK	KNOW	GREAT	LIVE	CANCER	COMPANY
FEEL	PRAY	MONTH	HELP	HUG	LIFE	ONCOLOGIST	HEALTH
WILL	KEEP	YEAR	FAMILY	GLAD	SURVIVOR	CAT	JOB
CAN	US	STAGE	MOTHER	THANK	DIAGNOSE	SPOT	CARE
GO	WELL	DIAGNOSE	WANT	WONDER	FIGHT	MRI	MEDIC
THINK	HUG	CANCER	CARE	ENJOY	CAN	BIOSPI	CANCER
HELP	THOUGHT	SURGERY	HOME	HOPE	SURVIVE	DOCTOR	MONEY
KEEP	WISH	FIRST	FATHER	CONGRATUL	POSIT	DOE	COST
JUST	SORRY	ROUND	TRY	CONGRAT	CURE	LUNG	COVER
WORK	BEST	LYMPH	CAREGIVER	CELEBRATION	DOCTOR	CHECK	EMPLOY
THANK	PLEASE	DAY	BROTHER	LOVE	DISEASE	TISSUE	SOCIAL
TREATMENT	LUCK	AUGUST	SISTER	WELL	SMOKE	LIGHT	FINANCIAL
FIND	GLAD	COMPLETE	FRIEND	BLESS	DIE	CHEST	SERVICE
TELL	GOD	JUNE	SUPPORT	BEST	BEAT	REPORT	MEDICARE
LOOK	CARE	2008	LIFE	AWESOME	HOPE	BODY	DOE
ALWAYS	BLESS	MAY	WIFE	HOLIDAY	BATTLE	BLOOD	PLAN
HARD	LOVE	TUMOR	HEART	NICE	SCARE	RADIOLOGIST	COVERAGE
SAID	RECOVERY	OLD	DAUGHTER	GOD	FEAR	SCAR	PROGRAM
MAKE	WORRY	TIME	PARENT	EXCITED	TREATMENT	RAY	BENEFIT
START	FINE	DEC	ALONE	FINISH	ENCOURAGE	LIVER	EXPENSE
COME	SCARE	LONG	COMFORT	JOY	DEATH	LYMPH	BILL
DONE	STRONG	SINCE	SICK	INSPIRE	LUNG	NODE	AMERICAN

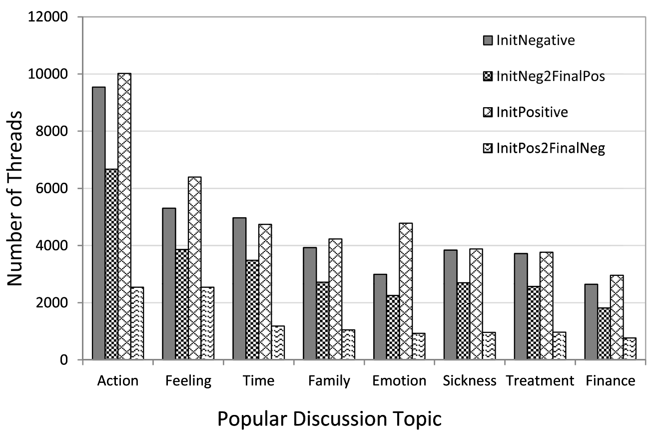


Fig. 5. Sentiment change as a function of popular discussion topic.

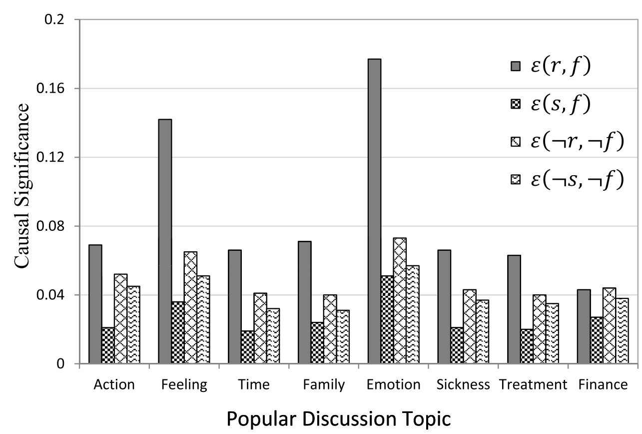


Fig. 6. Causal significance according to popular discussion topic.

temporal causality analysis of CSN sentiment dynamics for threads assigned to each discussion topic.

Table VII shows the main topics[5] that are discussed among participants in CSN data set (e.g., topic Emotion is discussed in 7777 out of 22 854 threads). Fig. 5 shows the sentiment change of the thread originators from an initially negative sentiment to a (finally) positive sentiment (InitNeg2FinalPos) and from an initially positive sentiment to (finally) negative sentiment (InitPos2FinalNeg) for each topic. Let r_{n2p} = (#InitNeg2FinalPos/#InitNeg) and r_{p2n} = (#InitPos2FinalNeg/#InitPos). We find that for each topic, $r_{n2p} > 68\%$ and $r_{p2n} > 25\%$. We proceed, for each topic, to examine the potential causes of f and $\neg f$ and find that $\{r, s\}$ and $\{\neg r, \neg s\}$ are potential causes of f and $\neg f$, respectively (similar to the ones in Section IV-A). When we repeat our analysis of causal significance for each topic (see Fig. 6), we find that r causally influences f and $\{\neg r, \neg s\}$ causally affect $\neg f$ (consistent with the results of our analysis based on the entire data set).

One interesting finding here is that the causal significance values in discussion threads relating to Feeling and Emotion topics are higher than those in the discussion threads relating to other topics. This finding is consistent with previous findings of [32] that discussion threads that involve Emotion or Feeling are more likely to be associated with change in sentiment than discussion threads that involve other topics. More importantly, it sheds light on the differences in causal significance of the factors contributing to sentiment change based on discussion topic.

[5]In each column, we present words which have the highest contribution to the corresponding topic, which is a distribution over a set of words (see [33] for more details). The number of topics given to LDA is 100.

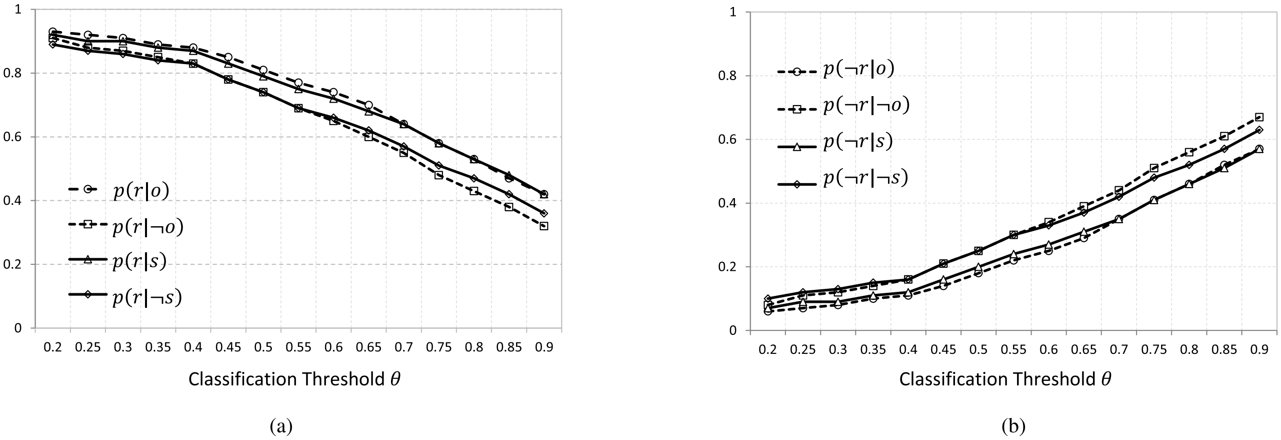


Fig. 7. Probability of reply sentiment conditioned on the post sentiment. (a) Positive reply transition. (b) Negative reply transition.

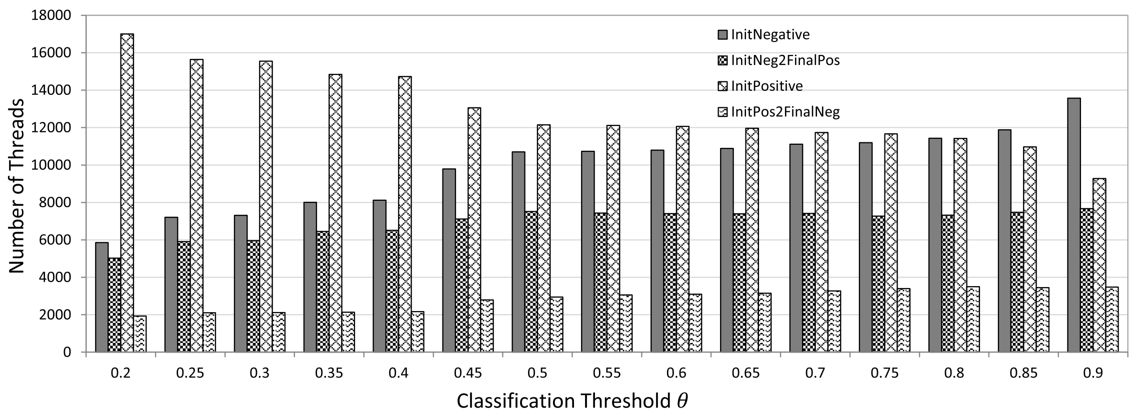


Fig. 8. Dynamic sentiment as a function of classification threshold θ.

V. Sensitivity of Inferred Temporal Causality of CSN Sentiment Dynamics With Respect to the Choice of Classification Threshold

Our preceding analysis of temporal causality of CSN sentiment dynamics relies on sentiment classification obtained by using a sentiment classifier. Recall that the output of the sentiment classifier depends on the choice of the classification threshold θ used to classify posts into positive and negative categories. Hence, we examine the sensitivity of the results of our analysis of temporal causality of CSN sentiment dynamics to the choice of θ. We choose a broad range of values for θ (0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90), the classification threshold used by the sentiment classifier and repeat our analysis of temporal causality of CSN sentiment dynamics for each choice of θ.

We first investigate how the choice of θ impacts the results of analysis of the sentiment of the replies as a function of the sentiment of the thread originator's posts. Fig. 7 shows us the positive and negative replies from repliers when they receive positive or negative posts from thread originators. Not surprisingly, as we increase θ, the probability of a positive (or negative) reply conditioned on the sentiment of thread originator's post [i.e., $p(r|x)$ and $p(\neg r|x)$ where $x \in \{o, \neg o, s, \neg s\}$] decreases (or increases). However, for a broad range of choices of θ in the neighborhood of $\theta = 0.5$, we find that repliers tend to express positive sentiment regardless the sentiment of the thread originators. In other words, members of CSN try to offer positive social support to others who seek support.

We examine the sentiment dynamics of the thread originators as a function of θ. Fig. 8 shows the dynamics of sentiment change of the thread originators (i.e., InitNeg2FinalPos and InitPos2FinalNeg). As θ increases, r_{n2p} (r_{p2n}) decreases (and increases). Specifically, r_{n2p} decreases from 0.86 to 0.57, and r_{p2n} increases from 0.11 to 0.37. This is not surprising, since the larger the value of θ, the lesser the number of posts classified as expressing positive sentiment.

The results in Fig. 8 show that the assessment of sentiment dynamics is fairly stable in the neighborhood of $\theta = 0.5$. We then investigate the potential causes of these sentiment dynamics. For each of the choice of θ, we construct a probabilistic Kripke structure that models the sentiment dynamics of a thread and repeat the analysis from Section IV. Table VIII shows the prima facie causes on thread originator's final sentiment for different choices of θ. In a majority of the

TABLE VIII
PRIMA FACIE CAUSES ACCORDING TO THE CLASSIFICATION THRESHOLD θ

θ	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90
f	r, s	r, s	r, s	r, s	r, s	r, s	r, s	r, s	r, s	r, s	r, s	r, s	r, s	r	r
$\neg f$	$\neg s$	$\neg r$	$\neg o, \neg r$	$\neg r$	$\neg r, \neg s$	$\neg r$	$\neg o, \neg r, \neg s$	$\neg r, \neg s$	$\neg r, \neg s$	$\neg r, \neg s$	$\neg r, \neg s$	$\neg r, \neg s$	$\neg r$	$\neg r, \neg s$	$\neg r$

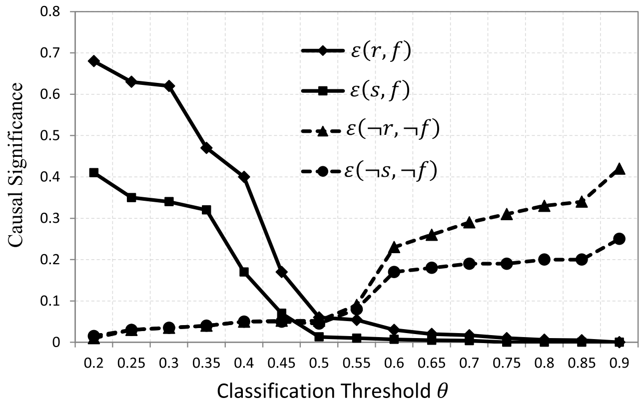


Fig. 9. Causal significance as a function of the classification threshold θ.

cases, $\{r, s\}$ and $\{\neg r, \neg s\}$ appear to be prima facie causes of f and $\neg f$, respectively.

Next, we examine the causal significance as a function of θ. Fig. 9 shows the causal significance of prima facie causes of the effects f and $\neg f$. Fig. 8 shows that when θ is small, there are many more positive sentiment classifications and much fewer negative sentiment classifications. The causal significance results in Fig. 9 show that the sentiment classifier's bias in favor of positive classification (i.e., low classification threshold θ) will exaggerate the role of positive replies and self-replies as causes of the final positive sentiment of the thread originator. On the other hand, the sentiment classifier's bias in favor of negative classification (i.e., high classification threshold θ) exaggerates the role of negative replies and self-replies as causes of the final negative sentiment of the thread originator. Fig. 9 also shows that the extents of causal effects are moderate when θ is in the neighborhood of 0.5 (i.e., 0.45–0.55) as compared with the extents of causal effects with respect to other choices on θ. When we applied the paired t-test on the causal strength of the prima facie causes of the effect f, we found that the causal significance of r on f is significantly higher ($p < 0.005$) than the causal significance of s on f. Likewise, causal significance of $\neg r$ on $\neg f$ is significantly higher ($p < 0.005$) than the causal significance of $\neg s$ on $\neg f$. Hence, we conclude that r and $\neg r$ are likely the genuine causes of f and $\neg f$, respectively.

VI. INCORPORATING SENTIMENT CLASSIFICATION ERROR RATES INTO THE ANALYSIS OF TEMPORAL CAUSALITY OF CSN SENTIMENT DYNAMICS

The preceding analysis of temporal causality of CSN sentiment dynamics uses a probabilistic Kripke structure [27] to represent and reason about probabilistic transitions between the sentiment states of posts in a CSN. However, since the true sentiments of the posts are unknown, we use a necessarily imperfect sentiment classifier to predict the sentiment states and use the predicted states as if they were the true states in constructing the probabilistic Kripke structure. Now we proceed to incorporate the error rates of the sentiment classifier directly into the analysis of CSN sentiment dynamics. Recall that positive predictive value (precision) PPV = (TP/(TP + FP)) and negative predictive value NPV = (TN/(TN + FN)), where TP, TN, FP, and FN are the numbers of true positive, true negative, false positive, and false negative labels (respectively) assigned by a two-class classifier (in our case, sentiment classifier).

Given a sentiment sequence $\mathbf{z} = z_1 z_2 \ldots z_n$, we define an indicator function $y(\mathbf{z}) = y_1 y_2 \ldots y_n \in \{0, 1\}^n$ where $y_i = 1$ if and only if $z_i \in \{o, r, s, f\}$; $y_i = 0$, otherwise. Let $N(\mathbf{z})$ be a number of occurrences of a sentiment sequence $\mathbf{z}$ in $\mathcal{D}$, the multiset of predicted sentiment sequences. Since the classifier is imperfect, $N(\mathbf{z})$ can be different from the true frequency of $\mathbf{z}$, the number of occurrences of $\mathbf{z}$ that we would observe if the sentiment labels were obtained by using a perfect sentiment classifier. Let $\widehat{M}(\mathbf{z})$ be an estimate of the true frequency of $\mathbf{z}$ given that the observed frequency (based on the use of an imperfect sentiment classifier) is $N(\mathbf{z})$. Let α = PPV and β = NPV of the sentiment classifier. Then, we can show that

$$\widehat{M}(\mathbf{z}) = N(\mathbf{z}) \prod_{i=1}^{n} (\alpha^{y_i} \beta^{1-y_i}) + \sum_{\mathbf{z}' : \mathbf{z}' \neq \mathbf{z} \wedge |\mathbf{z}'| = |\mathbf{z}|} N(\mathbf{z}') \prod_{i=1}^{n} (1-\alpha)^{y'_i} (1-\beta)^{1-y'_i} \quad (3)$$

where $\mathbf{z}'$ is a sentiment sequence, such that $|\mathbf{z}'| = |\mathbf{z}|$ (i.e., the same length) and $\mathbf{z}' \neq \mathbf{z}$; $y(\mathbf{z}') = y'_1 y'_2 \ldots y'_n \in \{0, 1\}^n$. We can prove formula (3) using induction.

1) Let $n = 1$, $\mathbf{z} = z_1$, and $y(\mathbf{z}) = y_1 \in \{0, 1\}$. Without loss of generality, let z_1 be a positive sentiment, in which case $y_1 = 1$ and $N(\mathbf{z})$ is the number of posts that are classified as expressing a positive sentiment. Let $\mathbf{z}' = z'_1$ and z'_1 be a negative sentiment, then $\mathbf{y}(\mathbf{z}') = y'_1$ and $y'_1 = 0$ and $N(\mathbf{z}')$ is the number of posts that are classified as expressing a negative sentiment. Then, the estimated number of posts, which truly express a positive sentiment, is given by

$$\begin{aligned} \widehat{M}(\mathbf{z}) &= N(\mathbf{z}) \times \alpha + (1-\beta) N(\mathbf{z}') \\ &= N(\mathbf{z}) \times \alpha^{y_1} \beta^{1-y_1} + N(\mathbf{z}') \\ &\quad \times (1-\alpha)^{y'_1} (1-\beta)^{1-y'_1}. \end{aligned} \quad (4)$$

TABLE IX
SENTIMENT CHANGE CALCULATIONS WITH SENTIMENT CLASSIFIER PPV $= \alpha$ AND NPV $= \beta$

	$\#[\neg o \to f] = M_1$	$\#[\neg o \to \neg f] = M_2$	$\#[o \to f] = M_3$	$\#[o \to \neg f] = M_4$
$\#[\neg o \to f]$	$M_1 \times \beta \times \alpha$	$M_2 \times \beta \times (1-\beta)$	$M_3 \times (1-\alpha) \times \alpha$	$M_4 \times (1-\alpha) \times (1-\beta)$
$\#[\neg o \to \neg f]$	$M_1 \times \beta \times (1-\alpha)$	$M_2 \times \beta \times \beta$	$M_3 \times (1-\alpha) \times (1-\alpha)$	$M_4 \times (1-\alpha) \times \beta$
$\#[o \to f]$	$M_1 \times (1-\beta) \times \alpha$	$M_2 \times (1-\beta) \times (1-\beta)$	$M_3 \times \alpha \times \alpha$	$M_4 \times \alpha \times (1-\beta)$
$\#[o \to \neg f]$	$M_1 \times (1-\beta) \times (1-\alpha)$	$M_2 \times (1-\beta) \times \beta$	$M_3 \times \alpha \times (1-\alpha)$	$M_4 \times \alpha \times \beta$

TABLE X
SENTIMENT CHANGE IN CSN (SENTIMENT CLASSIFIER PPV $= 0.84$ AND NPV $= 0.68$)

$\#[o]$	$\#[\neg o]$	$\#[\neg o \to f]$	$\#[o \to \neg f]$
13630	9224	6373	4000

2) Let $n = k$, then $\mathbf{z} = z_1 z_2 \ldots z_k$ and $y(\mathbf{z}) \in \{0,1\}^k$. Then, it follows (by induction on n) that:

$$\widehat{M}(\mathbf{z}) = N(\mathbf{z}) \prod_{i=1}^{k} (\alpha^{y_i} \beta^{1-y_i}) + \sum_{\mathbf{z}' : \mathbf{z}' \neq \mathbf{z} \wedge |\mathbf{z}'| = |\mathbf{z}|} N(\mathbf{z}') \prod_{i=1}^{k} (1-\alpha)^{y'_i} (1-\beta)^{1-y'_i} \tag{5}$$

where $\mathbf{z}'$ is a sentiment sequence, such that $|\mathbf{z}'| = |\mathbf{z}|$ and $\mathbf{z}' \neq \mathbf{z}$; $y(\mathbf{z}') = y'_1 y'_2 \ldots y'_k \in \{0,1\}^k$.

Hence, we can incorporate α and β into the estimation of MM transition probability as follows:

$$\hat{p}(X_i = \sigma | w) = \left[\frac{1 + \widehat{M}(w\sigma)}{|\mathcal{X}| + \sum_{\sigma' \in \mathcal{X}} \widehat{M}(w\sigma')} \right]_{\sigma \in \mathcal{X}}. \tag{6}$$

Next, we incorporate α and β into the causal significance calculation by modifying the original formula [29], which assumed perfect knowledge of the states. The contribution of a prima facie cause c to the change in probability of e can be written as follows:

$$\varepsilon_x(c, e) = \frac{\widehat{M}(cxe) + \widehat{M}(xce)}{\widehat{M}(cx) + \widehat{M}(xc)} - \frac{\widehat{M}(\neg cxe) + \widehat{M}(x\neg ce)}{\widehat{M}(\neg cx) + \widehat{M}(x\neg c)}. \tag{7}$$

By using formula (7) in the formula $\varepsilon(c, e) = (\sum_{x \in X \setminus c} \varepsilon_x(c, e) / |X \setminus c|)$, we obtain the causal significance with the incorporated error rates α and β. Our Adaboost classifier has $\alpha = 0.84$ and $\beta = 0.68$ (as estimated using 10-fold cross-validation). Based on these values of α and β, the estimated values of $\#[o]$ and $\#[\neg o]$ are $12147 \times 0.84 + (1 - 0.68) \times 10707 \approx 13630$ and $10707 \times 0.68 + (1 - 0.84) \times 12147 \approx 9224$, respectively. The calculations of the estimates of $\#[\neg o \to f]$ and $\#[o \to \neg f]$ transitions based on α and β are summarized in Table IX. Here, M_1, M_2, M_3, and M_4 are the corresponding estimates from Table III (which correspond to the assumption that $\alpha = 1$ and $\beta = 1$). From Table III, we have $M_1 = 7512$, $M_2 = 3195$, $M_3 = 9199$, and $M_4 = 2948$. Sentiment dynamics results after incorporating α and β are shown in Table X. The results show that approximately 69.1% of thread originators with an initially negative sentiment end up with a positive sentiment and 29.3% of thread originators with an initially positive sentiment end up with a negative sentiment at the end. These results are consistent with the results presented in Section IV (which correspond to the assumption that $\alpha = 1$ and $\beta = 1$, i.e., we have a perfect sentiment classifier).

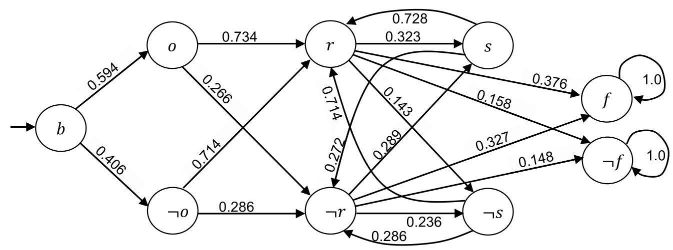


Fig. 10. Probabilistic Kripke structure for CSN sentiment change (sentiment classifier PPV $= 0.84$ and NPV $= 0.68$).

TABLE XI
CAUSAL SIGNIFICANCE ESTIMATES (SENTIMENT CLASSIFIER PPV $= 0.84$ AND NPV $= 0.68$)

$\varepsilon_{avg}(r, f)$	$\varepsilon_{avg}(s, f)$	$\varepsilon_{avg}(\neg r, \neg f)$	$\varepsilon_{avg}(\neg s, \neg f)$
0.087	0.005	0.062	0.053

Next, we construct the probabilistic Kripke structure and investigate whether there exist prima facie causes for the final sentiment of the thread originators. Fig. 10 shows the probabilistic Kripke structure after we apply formula (6) to estimate the transition probabilities between sentiment states. The resulting probabilistic Kripke structure shows that from any state of the thread originator, i.e., $\{o, \neg o, s, \neg s\}$, there is a probability greater than 71% that it will transit to the state r. In other words, members of CSN try to offer positive social support to others who seek support (a result that is consistent with that presented in Section IV). As in Section IV, we find that $\{r, s\}$ and $\{\neg r, \neg s\}$ are prima facie causes of f and $\neg f$, respectively.

We proceed to evaluate the significance of the prima facie causes of f and $\neg f$ using the modified causal significance formula (7), which incorporates PPV and NPV of the sentiment classifier. The results of our analysis are shown in Table XI, where r and $\neg r$, $\neg s$ are causally significant with respect to f and $\neg f$, respectively.

VII. ROBUSTNESS OF INFERRED TEMPORAL CAUSALITY OF CSN SENTIMENT DYNAMICS WITH RESPECT TO THE CHOICE OF SENTIMENT CLASSIFIERS

The preceding sections (Sections V and VI) examined the robustness of the inferred temporal causality based on the

TABLE XII
CAUSAL SIGNIFICANCE OF SENTIMENT DYNAMICS WHEN USING CLASSIFICATION RESULTS OF SVM (C_1) AND LOGISTIC REGRESSION (C_2). C_1' AND C_2' ARE CLASSIFICATION RESULTS WHEN INCORPORATING PPV AND NPV OF SVM AND LOGISTIC REGRESSION, RESPECTIVELY

	$\varepsilon(r,f)$	$\varepsilon(s,f)$	$\varepsilon(\neg r,\neg f)$	$\varepsilon(\neg s,\neg f)$
C_1	0.106	0.017	0.045	0.033
C_1'	0.086	0.0005	0.075	0.063
C_2	0.191	0.057	0.051	0.04
C_2'	0.124	0.026	0.065	0.048

sentiment labels assigned to posts by a specific classifier (i.e., Adaboost, described in Section III-A). In this section, we further explore the robustness of the inferred temporal causality sentiment dynamics in CSN with respect to the specific choice of the sentiment classifier. Specifically, we repeat our analyses using two different sentiment classifiers trained on the same training data, as described in Section III-A: SVM[6] and logistic regression[7] to label the entire set of CSN posts. The resulting SVM classifier had a PPV ($\alpha = 0.8$) and NPV ($\beta = 0.7$); and the logistic regression classifier had PPV ($\alpha = 0.8$) and NPV ($\beta = 0.67$), in both cases estimated using 10-fold cross-validation on the training data. We proceed to identify prima facie and genuine causes of sentiment dynamics of the thread originators. We also repeated the analysis using the modified procedure for causal significance assessment that incorporates the error rates of the classifiers into the analysis as described in Section VI. The results of these analyses are summarized in Table XII. We see that the causal significances of prima facie causes $\{r, s\}$ and $\{\neg r, \neg s\}$ on the effects f and $\neg f$ are consistent with the findings obtained using the Adaboost classifier: r and $\neg r$, $\neg s$ are causally significant with respect to f and $\neg f$, respectively.

VIII. CONCLUSION AND DISCUSSION

In this paper, we have introduced a framework to uncover the temporal causality of sentiment dynamics in the American Cancer Society's CSN. To the best of our knowledge, this paper is the first to uncover the factors that causally drive the sentiment dynamics in an OHC. We developed a sentiment classifier using machine learning on a training set of posts manually labeled for their sentiment (positive versus negative). We constructed a PCTL representation and a corresponding probabilistic Kripke structure to represent and reason about the transitions between sentiments of posts in a thread over time. We analyzed the probabilistic Kripke structure to identify the prima facie causes of sentiment change on the part of the thread originators in the CSN forum and their significance. Our main finding is that the positive sentiment of replies appears to causally influence the positive sentiment of the thread originator at the end of the thread; conversely, the negative sentiment of the replies appears to causally influence the negative sentiment of the thread originator at the end of the thread. Our methodology can be used to gain new insights that the designers, managers, and moderators of an online community, such as CSN, can utilize to facilitate and enhance the interactions so as to better meet the social support needs of the CSN participants. The methodology for analysis of temporal causality has broad applicability in a variety of settings where the dynamics of the underlying system can be modeled in terms of state variables that change in response to internal or external inputs.

Our methods, such as most existing computational and statistical approaches to causal inference from observational or experimental data, make several key assumptions that may or may not hold in practice.

1) *Causal Sufficiency* [15], [17], [34] that requires that all shared causes of the variables over which we are performing causal inference are captured in the data; missing variables can lead to spurious findings. For example, if caffeine leads to sleep deprivation and to increase in one's heart rate, if we don't measure caffeine consumption, we might incorrectly infer causal relationships among its effects sleep deprivation, heart rate.
2) *Representativeness of Data* [34] that, roughly translated, is tantamount to assuming that the observed data reflect the true behavior of the system being studied. This assumption can be violated for a variety of reasons, e.g., selection bias, missing measurements, presence of multiple causal chains whose effects on a variable essentially cancel out (which corresponds to causation in the absence of correlation), and insufficient number of observations (leading to inaccurate estimates of probabilities).
3) *Stationarity* of the underlying process, which if violated, could mean that some of the relationships between causes and effects may change over time.

It should be further noted that the validity of our conclusions depends critically on the accuracy of sentiment labels (assigned by the annotators in the case of labeled training data, and the sentiment labels assigned to the rest of the posts by the classifier trained on the labeled training data). Hence, we examined the robustness of inferred temporal causality of CSN sentiment dynamics with respect to the choice of the: 1) classification threshold of the sentiment classifier and 2) choice of the specific sentiment classifier used. We also extend the basic framework for analysis of temporal causality of sentiment dynamics that incorporates uncertainty in sentiment classification, and hence the states of the probabilistic Kripke structure resulting from the use of an imperfect state transducer (in our case, sentiment classifier). Temporal causality analysis using the resulting modified probabilistic Kripke structure shows that the causal effects inferred using the modified Kripke structure are consistent with those obtained using the unmodified Kripke structure. It should be further noted that, at present, we have no ground truth regarding the actual sentiments of the CSN participants during the course of their online interactions. Hence, the conclusions

[6]Using the LibSVM implementation with default parameters, https://www.csie.ntu.edu.tw/ cjlin/libsvm/

[7]Using the default implementation in Weka, http://www.cs.waikato.ac.nz/ml/weka/

drawn from our analyses are best viewed as hypotheses to be validated by controlled experiments, and examination by experts.

Some promising directions for future research include: 1) exploring the causal effects of the discussion topic on the sentiment dynamics; 2) exploring the causal effects of (explicit as well as implicit) social relations among OHC participants on sentiment dynamics; 3) extending the framework to handle unobserved variables; 4) scaling up temporal causality analysis to very large state spaces; and 5) extending the framework to handle relational data [35], [36].

REFERENCES

[1] J. Ferlay *et al.*, "Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012," *Int. J. Cancer*, vol. 136, no. 5, pp. E359–E386, 2014.
[2] American Cancer Society. (2014). *Cancer Facts and Figures 2014*. [Online]. Available: http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2014/
[3] S. Fox and M. Duggan. (2013). *Health Online 2013*. [Online]. Available: http://www.pewinternet.org/files/old-media/Files/Reports/PIP_Health Online.pdf
[4] C. Dunkel-Schetter, "Social support and cancer: Findings based on patient interviews and their implications," *J. Social Issues*, vol. 40, no. 4, pp. 77–98, 1984.
[5] J. Preece, "Empathic communities: Balancing emotional and factual communication," *Interact. Comput.*, vol. 12, pp. 63–77, Sep. 1999.
[6] S. Rodgers and Q. Chen, "Internet community group participation: Psychosocial benefits for women with breast cancer," *J. Comput.-Mediated Commun.*, vol. 10, no. 4, pp. 00–00, Jul. 2005.
[7] C. E. Beaudoin and C.-C. Tao, "Modeling the impact of online cancer resources on supporters of cancer patients," *New Media Soc.*, vol. 10, no. 2, pp. 321–344, 2008.
[8] D. Maloney-Krichmar and J. Preece, "A multilevel analysis of sociability, usability, and community dynamics in an online health community," *ACM Trans. Comput.-Human Interact.*, vol. 12, no. 2, pp. 201–232, Jun. 2005.
[9] G. Bouma, J. M. Admiraal, E. G. E. de Vries, C. P. Schröder, A. M. E. Walenkamp, and A. K. L. Reyners, "Internet-based support programs to alleviate psychosocial and physical symptoms in cancer patients: A literature analysis," *Critical Rev. Oncol./Hematol.*, vol. 95, no. 1, pp. 26–37, 2015.
[10] K. Portier *et al.*, "Understanding topics and sentiment in an online cancer survivor community," *J. Nat. Cancer Inst. Monogr.*, vol. 47, pp. 195–198, Dec. 2013.
[11] B. Qiu *et al.*, "Get online support, feel better—Sentiment analysis and dynamics in an online cancer survivor community," in *Proc. SocialCom/PASSAT*, 2011, pp. 274–281.
[12] J. Huh, M. Yetisgen-Yildiz, and W. Pratt, "Text classification for assisting moderators in online health communities," *J. Biomed. Inform.*, vol. 46, no. 6, pp. 998–1005, Dec. 2013.
[13] X. Wang, K. Zhao, and N. Street, "Social support and user engagement in online health communities," in *Proc. Int. Conf. Smart Health*, 2014, pp. 97–110.
[14] N. Bui, J. Yen, and V. Honavar, "Temporal causality of social support in an online community for cancer survivors," in *Proc. 8th Int. Conf. Social Comput., Behavioral-Cultural Modeling, Predict.*, 2015, pp. 13–23.
[15] J. Pearl, *Causality: Models, Reasoning, and Inference*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
[16] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search*. Cambridge, MA, USA: MIT Press, 2000.
[17] J. Pearl, M. Glymour, and N. P. Jewell, *Causal Inference in Statistics: A Primer*. New York, NY, USA: Wiley, 2016.
[18] S. L. Morgan and C. Winship, *Counterfactuals and Causal Inference*. Cambridge, U.K.: Cambridge Univ. Press, 2014.
[19] G. W. Imbens and D. B. Rubin, *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge, U.K.: Cambridge Univ. Press, 2015.
[20] P. Dagum, A. Galper, and E. Horvitz, "Dynamic network models for forecasting," in *Proc. 8th Conf. Uncertainty Artif. Intell.*, 1992, pp. 41–48.
[21] N. Friedman, K. Murphy, and S. Russell, "Learning the structure of dynamic probabilistic networks," in *Proc. 14th Conf. Uncertainty Artif. Intell.*, 1998, pp. 139–147.
[22] Z. Ghahramani, "Learning dynamic Bayesian networks," in *Adaptive Processing of Sequences and Data Structures*. Berlin Germany: Springer, 1998, pp. 168–197.
[23] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica, J. Econometric Soc.*, vol. 37, no. 3, pp. 424–438, 1969.
[24] S. Kleinberg, *Causality, Probability, and Time*. Cambridge, U.K.: Cambridge Univ. Press, 2013.
[25] P. Suppes, *A Probabilistic Theory of Causality*. Amsterdam, The Netherlands: North Holland, 1970.
[26] A. Prior, *Past, Present, and Future*. Oxford, U.K.: Clarendon, 1967.
[27] E. M. Clarke, Jr., O. Grumberg, and D. A. Peled, *Model Checking*. Cambridge, MA, USA: MIT Press, 1999.
[28] H. Hansson and B. Jonsson, "A logic for reasoning about time and reliability," *Formal Aspects Comput.*, vol. 6, no. 5, pp. 102–111, 1994.
[29] S. Kleinberg and B. Mishra, "The temporal logic of causal structures," in *Proc. UAI*, Arlington, VA, USA, 2009, pp. 303–312.
[30] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas, "Sentiment in short strength detection informal text," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 61, no. 12, pp. 2544–2558, Dec. 2010.
[31] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proc. ACM KDD*, New York, NY, USA, 2004, pp. 168–177.
[32] P. Biyani, C. Caragea, P. Mitra, and J. Yen, "Identifying emotional and informational support in online health communities," in *Proc. COLING*, 2014, pp. 827–836.
[33] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.
[34] S. Kleinberg, *Why: A Guide to Finding and Using Causes*. Sebastopol, CA, USA: O'Reilly Media, Inc., 2015.
[35] S. Lee and V. Honavar, "On learning causal models from relational data," in *Proc. 30th AAAI Conf. Artif. Intell.*, Phoenix, AZ, USA, Feb. 2016, pp. 3263–3270.
[36] S. Lee and V. Honavar, "A characterization of Markov equivalence classes of relational causal models under path semantics," in *Proc. Conf. Uncertainty Artif. Intell.*, 2016, pp. 387–396.

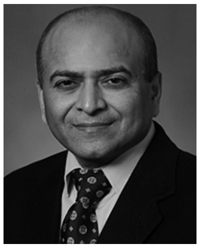
Ngot Bui received the B.Sc. degree in computer engineering from Vietnam National University, Ho Chi Minh City, Vietnam, in 2004, and the M.Sc. degree in computer engineering from Kyung Hee University, Seoul, South Korea, in 2008. He is currently pursuing the Ph.D. degree with the Artificial Intelligence Research Laboratory, College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA, USA.

His current research interests include machine learning, data mining, social network analysis, sentiment analysis, and bioinformatics.

John Yen (F'00) received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1980, the M.S. degree in computer science from the University of Santa Clara, Santa Clara, CA, USA, in 1982, and the Ph.D. degree in computer science from the University of California at Berkeley, Berkeley, CA, USA, in 1986.

He was a Research Scientist with the Information Sciences Institute, University of Southern California, Los Angeles, CA, USA, from 1986 to 1989. From 1990 to 2001, he served on the faculty of Computer Science, Texas A&M University, College Station, TX, USA. He is currently a University Professor and Director of Strategic Research Initiatives with the College of Information Sciences and Technology, The Pennsylvania State University (Penn State), University Park, PA, USA. He has authored over 200 refereed papers in journals and conference proceedings. His current research interests include artificial intelligence, knowledge representation, machine learning, large-scale multimodality data analytics, and computational modeling.

Dr. Yen is a Senior Member of the Association for the Advancement of Artificial Intelligence. He received an NSF Young Investigator Award in 1992, the IBM Faculty Award in 2012, and Penn State's University Graduate Teaching Award in 2015.



Vasant Honavar (SM'06) received the B.E. degree in electronics engineering from Bangalore University, Bangalore, India, the M.S. degree in electrical and computer engineering from Drexel University, Philadelphia, PA, USA, and the M.S. and Ph.D. degrees in computer science from the University of Wisconsin–Madison, Madison, WI, USA.

He was on the faculty of Computer Science, Iowa State University, Ames, IA, USA, from 1990 to 2013. From 2010 to 2013, he served as a Program Director of the Information and Intelligent Systems Division with the National Science Foundation (NSF), where he led the Big Data Program, and contributed to the Smart and Connected Health, Information Integration and Informatics, and Expeditions in Computing programs. He currently holds the Edward Frymoyer Endowed Professorship in Information Sciences and Technology, and serves as a Professor of Computer Science, Bioinformatics and Genomics, and Neuroscience with The Pennsylvania State University, University Park, PA, USA. He also directs the Artificial Intelligence Research Laboratory and the Center for Big Data Analytics and Discovery Informatics, co-directs the NIH-Funded Biomedical Big Data to Knowledge Pre-Doctoral Training Program, and serves as an Associate Director of the Penn State Institute for CyberScience. His work (documented in over 250 peer-reviewed publications, with over 10,300 citations during 1990 to 2016) has resulted in foundational contributions in scalable approaches to learning predictive models from very large, richly structured data, including tabular, sequence, network, relational, and time series data; eliciting causal information from multiple sources of observational and experimental data; selective sharing of knowledge across disparate knowledge bases, including novel approaches to secrecy-preserving query answering; representing and reasoning about preferences; composing complex software services from components; and applications in bioinformatics and computational molecular and systems biology, including characterization, analysis, and prediction of sequence and structural correlates of protein–protein, and protein–RNA interfaces and interactions, comparative analysis of biological networks (network alignment). His current research interests include artificial intelligence, machine learning, bioinformatics, big data analytics, discovery informatics, brain informatics, knowledge representation, causal inference, and health informatics.

Dr. Honavar is a Senior Member of the Association for the Advancement of Artificial Intelligence and the Association for Computing Machinery. He has received many awards and honors during his career, including the NSF Director's Award for Superior Accomplishment in 2013 for his leadership of the NSF Big Data Program, the NSF Director's Award for Collaborative Integration in 2011, the Iowa Board of Regents Award for Faculty Excellence in 2007, the Iowa State University College of Liberal Arts and Sciences Award for Career Excellence in Research in 2008, the Iowa State University Margaret Ellen White Graduate Faculty Award in 2011, and the Pennsylvania State University College of Information Sciences and Technology Research Excellence Award in 2016. He serves on the Computing Research Association Computing Community Consortium Council, the Electorate Nominating Committee of the Section on Information, Computing, and Communication of the American Association for the Advancement of Science, and on the Executive Committee of the North East Big Data Hub. He served as a General Co-Chair of the IEEE Conference on Big Data in 2014 and on the editorial boards of several journals and program committees of major conferences in artificial intelligence, machine learning, and bioinformatics.