Social Media Data Assisted Inference with Application to Stock Prediction

Hao He, Arun Subramanian, Sora Choi, Pramod K. Varshney
Dept. of Electrical Engineering & Computer Science
Syracuse University, Syracuse, NY 13244 U.S.A.

Thyagaraju Damarla
US Army Research Lab
2800 Powder Mill Rd., Adelphi, MD 20783 U.S.A.

***Abstract*—The access to the massive amount of social media data provides a unique opportunity to the signal processing community for extracting information that can be used to infer about unfolding events. It is desirable to investigate the convergence of sensor networks and social media in facilitating the data-to-decision making process and study how the two systems can complement each other for enhanced situational awareness. In this paper, we propose a copula-based joint characterization of multiple dependent time series from sensors and social media. As a proof-of-concept, this model is applied to the fusion of Google Trends (GT) data and stock price data of Apple Inc. for prediction, where the stock data serves as a surrogate for sensor data. Superior prediction performance is demonstrated, by taking the non-linear dependence among social media data and sensor data into consideration.**

I. INTRODUCTION

Sensor networks provide information about various aspects of the real world and have become an integral part of various systems used in daily lives. The problem of inferring about events of interest by fusing data from multiple heterogeneous sensors has a wide variety of applications. The inference tasks could consist of detecting an activity of interest or estimating some parameters, such as locations or tracks, which provide actionable intelligence and/or improved situational awareness.

Social media, facilitated by the growth of social networks, provides an easily accessible platform for users to share information and has resulted in the generation of unprecedented amounts of social media data that can be recorded and even monitored (such as, wall posts, clicks etc). This trend is likely to continue with exponentially more content in the future. This massive amount of social media data can be used by the signal processing community for extracting information about unfolding events. This is expected to be beneficial in the military as well as civilian domains. A number of works have been published regarding the use of social media data for understanding real world phenomena. In particular, social media data have been successfully applied to: prediction of earthquakes [1], forecast box office revenues [2], truth discovery [3], prediction of election results [4], stock prediction [5], and automatic crime prediction [6].

For inference with both traditional sensors and social media, with respect to the information content of the signals, information sources exhibit heterogeneity that can arise from a wide variety of causes. As a consequence of heterogeneity, the quality and quantity of information provided by each modality, including human intelligence, vary with each source. For a group of heterogeneous information sources observing the same phenomenon, local observations are statistically dependent, and yet provide different characterizations of the phenomenon under observation. Such diversity in the sensing process not only leads to enhanced inference quality, but also improves fault tolerance capabilities, so that the decision making capability of the system does not become impaired completely due to the failure of one modality. However, an accurate characterization of the intermodal dependence and development of algorithms to jointly process and fuse heterogeneous data are necessary for making reliable system-wide inference.

In this paper, we develop new techniques for inference, using data from diverse information sources, including both social media and sensor networks. The inference in traditional heterogeneous sensor networks has been investigated in [7], [8] using copula theory, but assuming that the observations are temporally independent. In the networks of sensors and social media, temporal dependence may exist, along with intermodal dependence. We carry out a copula-based characterization of multivariate time series, in which the marginal conditional distributions are modeled first; then copula theory is applied to approximate the dependence among the residual terms of the marginals. As a proof-of-concept, the copula-based model is used in the Google Trends (GT) data assisted stock prediction, and superior performance, in terms of Mean Square Error (MSE), is shown.

A. Background: copula theory

Copulas are parametric functions that couple univariate marginal distributions to a valid multivariate distribution. They explicitly model the dependence among random variables, which may have arbitrary marginal distributions. Copula theory is an outcome of the work on probabilistic metric spaces [9] and a copula was initially defined, on the unit hypercube, as a joint probability distribution for uniform marginals. Their application to statistical inference is possible largely due to Sklar's theorem, which is stated below without proof [10].

Theorem 1 (Sklar's Thoerem)**.** *Consider an N-dimensional distribution function F with marginal distribution functions*

Research was sponsored by ARO grant W911-NF-13-2-0040.

$F_1, \dots, F_N$. *Then there exists a copula C, such that for all $x_1, \dots, x_N$ in $[-\infty, \infty]$*

$$F(x_1, \dots, x_N) = C(F_1(x_1), \dots, F_N(x_N)) \tag{1}$$

If F_n is continuous for $1 \le n \le N$, then C is unique.

Note that the arguments of C in (1) are uniformly distributed random variables. As a direct consequence of Sklar's Theorem, for continuous distributions, the joint probability density function (PDF) is obtained by differentiating both sides of (1),

$$f(x_1, \dots, x_N) = \left(\prod_{n=1}^{N} f_n(x_n) \right) c(F_1(x_1), \dots, F_N(x_N)) \tag{2}$$

where, c is termed the copula density and is given by,

$$c(\mathbf{u}) = \frac{\partial^N C(u_1, \dots, u_N)}{\partial u_1, \dots, \partial u_N} \tag{3}$$

where, $u_n = F_n(x_n)$. Eq. (2) implies that we may link together any univariate distributions, of any type (not necessarily from the same family), with any copula and we will have defined a valid multivariate distribution. Several copula functions are defined in the literature, and are constructed to characterize different types of dependence [10], of which the elliptical and Archimedean copulas are widely used. While not explicitly specified in (1) and (2), copula functions contain a *dependence parameter* that quantifies the amount of dependence between the N random variables. We denote the dependence parameter as $\boldsymbol{\theta}_d$, which, in general, may be a scalar, a vector or a matrix.

The remainder of this paper is organized as follows. In Section II, the problem of inference with sensor data and social media data is formulated. In Section III, copula-based joint characterization of heterogeneous N-variate time series is presented. The example of stock prediction with the assistance of GT data is shown in Section IV.

II. Problem Formulation

We consider a network consisting of s traditional sensors and a set of "human sensors" providing a massive amount of accessible social media data, as shown in Figure 1. Traditional sensors take successive measurements about the phenomenon of interest over a time interval $l = 1, \dots, L$, while social media data about the same event is successively collected over the same time interval. Informative features, such as volume and sentiment, are, then, extracted from the social media data for inference purposes.

We use $x_{n,l}, n \in \{1, \dots, s\}$ to denote the observation of traditional sensor n and we use $x_{n,l}, n \in \{s+1, \dots, N\}$ to denote the extracted feature n of the social media data, at time instant l. The inference task is conducted at the fusion center based on the N-dimensional time series $\{x_{1,l}, \dots, x_{N,l}\}_{l=1}^{L}$ obtained from the sensors and social media. A statistical model is needed to characterize the vector stochastic process before any inference task, such as detection, estimation, and prediction, can be conducted. Since sensor data and social media data provide noisy observations of the same phenomenon, they are dependent across modalities. As discussed in Section I,

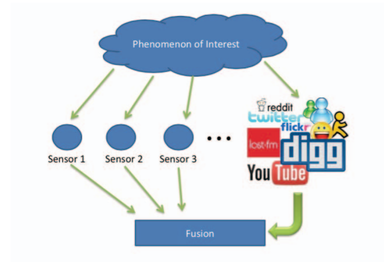

Fig. 1. The network of traditional sensors and social media.

observations from different sources, especially when including both sensors and social media, are very likely to be heterogeneous where the heterogeneity under the assumption that the N-dimensional random vector are i.i.d. in time is defined by Iyengar et al. in [7] as follows

Definition 1 (S. Iyengar [7])**.** *A random vector $\{X_1, \dots, X_N\}$ governing the joint statistics of N-variate data set is termed heterogeneous if the marginals X_n are non-identically distributed.*

Considering temporal dependence, which exists in most consecutive measurements but neglected for simplification, we extend the above definition of heterogeneity to a vector stochastic process as follows

Definition 2. *A vector stochastic process $\{X_{1,l}, \dots, X_{N,l}\}_{l=1}^{L}$ with dimension N is heterogeneous if the conditional marginal distributions $f_n(x_{n,l}|\,\mathbf{x}_n^{l-1})$ are non-identical, where $\mathbf{x}_n^{l-1} := \{x_{n,l}\}_{l=1}^{l-1}$.*

In all, there are two main properties of social data and sensor data that need to be captured by the statistical model:

- Two-dimensional dependence: temporal and intermodal dependence.
- Heterogeneity: disparate conditional marginal distributions.

To accommodate the above requirement of the statistical characterization of sensor data and social media data, in this paper we formulate a conditional copula-based approach to approximate the joint distribution. The following assumption is made about the observations.

Assumption 1. *The temporal dynamics of each time series does not depend on any other time series, i.e.,*

$$f_n(x_{n,l}|\mathcal{I}^{l-1}) = f_n(x_{n,l}|\,\mathbf{x}_n^{l-1}), \forall n = 1, \dots, N$$

where $\mathcal{I}^{l-1} := \{x_{1,l}, \dots, x_{N,l}\}_{l=1}^{l-1}$ denotes the entire observation set before time instant l.

III. Copula-based Multivariate Dynamic Models

Before modeling the dependence structure between the N time series using copulas, we need to model their conditional marginal distributions.

TABLE I
BEHAVIORS OF THEORETICAL ACF AND PACF

Model	ACF	PACF
AR(P)	exponential decay and/or damped sinusoid	cuts off after lag P
MA(Q)	cuts off after lag Q	exponential decay and/or damped sinusoid
ARMA(P,Q)	exponential decay and/or damped sinusoid	exponential decay and/or damped sinusoid

A. Conditional marginal distributions

We focus on the first two moments of the marginal conditional distributions. Assuming that the first two moments exist, we apply the following structure, which is commonly used to represent time series with time-varying conditional mean and time-varying conditional variance [11], to model each marginal

$$x_{n,l} = \mu_n(\mathbf{x}_n^{l-1}|\boldsymbol{\theta}_n) + \sigma_n(\mathbf{x}_n^{l-1}|\boldsymbol{\theta}_n)\epsilon_{n,l} \tag{4}$$

where $\boldsymbol{\theta}_n$ represents the set of unknown parameters and $\epsilon_{n,1}, \epsilon_{n,2}, \ldots, \epsilon_{n,L}$ is a sequence of i.i.d. random variables with zero mean and unit variance, but without a specified PDF. The model in (4) allows each time series to have a time-varying conditional mean and time-varying conditional variance.

The Auto Regressive Moving Average ARMA model [12] provides a description of the conditional mean part of Eq. (4). An ARMA(P, Q) model, where P is the order of the autoregressive part and Q is the order of the moving average part, is written as

$$x_{n,l} = \alpha_0 + \sum_{i=1}^{P} \alpha_i x_{n,l-i} + \sum_{j=1}^{Q} \beta_j z_{n,l-j} + z_{n,l} \tag{5}$$

where $z_{n,l}$ are the error terms that are generally assumed to be i.i.d random variables and with zero mean.

The Generalized autoregressive conditional heteroskedasticity (GARCH) model [13], [14] characterizes the conditional variance of the error term $z_{n,l} = \sigma_{n,l}\epsilon_{n,l}$ by imposing alternative parameters to capture serial dependence on the past sequence of observations as

$$\sigma_{n,l}^2 = a_0 + \sum_{i=1}^{M} a_i \sigma_{n,l-i}^2 + \sum_{j=1}^{N} b_j z_{n,l-j}^2 \tag{6}$$

In order to roughly determine the orders of both AR and MA parts of the model, we consider the autocorrelation function (ACF) and partial autocorrelation function (PACF).[1] The theoretical behaviors of the ACF and PACF are summarized in Table I [12]. As can be seen in Eq. (6), if an ARMA model is assumed for the error variance $\sigma_{n,l}^2$, the model is a GARCH model. Thus, the approximate order of GARCH(M, N) can be obtained by checking the ACF and PACF of the squared residuals $z_{n,l}^2$.

As to determining the optimal orders of the ARMA and GARCH models, there are a variety of model selection criteria such as the Akaike information criterion (AIC) and Bayesian information criterion (BIC), which are measures of the relative quality of a statistical model for a given set of data. That is, given a collection of models for the data, AIC/BIC estimate the quality of each model, relative to the other models. They are formally defined as

$$\begin{aligned} AIC &= -2\log\mathcal{L} + 2k \\ BIC &= -2\log\mathcal{L} + k\ln L \end{aligned} \tag{7}$$

where $\mathcal{L}$ is the maximized value of the likelihood function for the model; k is the number of parameters in the model (i.e. k is the number of degrees of freedom) and L is the number of observations, or equivalently, the sample size.

After fitting each time series to the selected models, the parameters of the ARMA-GARCH models are obtained and the estimated standardized residuals are

$$\hat{\epsilon}_{n,l} = \frac{x_{n,l} - \mu_n(\mathbf{x}_n^{l-1}|\hat{\theta}_n)}{\sigma_n(\mathbf{x}_n^{l-1}|\hat{\theta}_n)} \tag{8}$$

where $\hat{\theta}_n$ is the estimate of the parameters of each time series, for all $n = 1, \ldots, N$.

While a scalar GARCH model is used to capture the time-varying variance of each individual time series in this paper, a parametric copula is used to model the intermodal dependence between different time series.

B. Estimation and inference for copula models

In this subsection, we introduce the inference process for copula-based multivariate models. The conditional copula is defined as the joint distribution of the probability integral transforms of the standardized residuals. We consider the parametric copula model and its parameters to be invariant with time.

An important benefit of using copulas to construct multivariate models is that the models used in the marginal distributions need not be of the same type as the model used for the copula. One exciting possibility that this allows is semi-parametric estimation of the marginal distributions, combined with parametric estimation of the copula. Such a model avoids the curse of dimensionality by only estimating the one-dimensional marginal distributions non-parametrically, and then estimating the copula parametrically. In the semi-parametric model, the marginal distributions are modeled non-parametrically, using distributions such as the empirical distribution function (EDF) and a parametric model is used for intermodal dependence. When copula c is used to characterize the dependence structure, its corresponding parameters are estimated using the MLE based approach, as follows

$$\hat{\boldsymbol{\theta}}_d = \arg\max_{\boldsymbol{\theta}_d} \sum_{l=1}^{L} \log c(\hat{u}_{1,l}, \ldots, \hat{u}_{N,l}|\boldsymbol{\theta}_d) \tag{9}$$

where

$$\hat{u}_{n,l} = \hat{F}_{\epsilon_n}(\epsilon_{n,l}), \quad \forall n = 1, \ldots, N \tag{10}$$

[1]PACF is the autocorrelation after adjusting for a common factor.

We consider the nonparametric estimate of the CDF F_n using the EDF [2]:

$$\hat{F}_{\epsilon_n}(\epsilon) = \frac{1}{L+1}\sum_{l=1}^{L} \mathbb{1}_{\{\hat{\epsilon}_{n,l} \leq \epsilon\}} \tag{11}$$

where $\mathbb{1}_{\{\cdot\}}$ is the indicator function and the estimated standardized residual $\hat{\epsilon}_{n,l}$ is given in Eq. (8).

When the true dependence structure is unknown, the best copula model needs to be selected from a finite set of predefined copula density functions $\mathcal{A} = \{c_1, \ldots, c_m\}$. There are several copula selection approaches that can be employed, such as AIC and BIC based approaches. We use an MLE based approach, since for fixed N, different copulas do not have much difference in parameter dimensions. Thus, the best copula c^* is selected as follows

$$c^* = \arg\max_{c \in \mathcal{A}} \sum_{l=1}^{L} \log c(\hat{u}_{1,l}, \ldots, \hat{u}_{N,l} | \hat{\boldsymbol{\theta}}_d) \tag{12}$$

where $\hat{\boldsymbol{\theta}}_d$ is obtained using MLE in Eq. (10).

We apply the copula-based characterization of multivariate time series to social media data assisted stock prediction as an application of our methodology in the next section.

IV. Stock prediction with Google Trends data

The analysis and forecasting of stock market behavior has been a focus of academics and practitioners alike. A model that accounts for investor attention can provide a better explanation of the stock behavior and can also contain useful information to forecast stock market. Google Trends (GT) serves as a good "attention indicator" that is measured through social media. GT is a public web facility of Google Inc., based on Google Search, that shows how often a particular search-term is entered. When one searches for a term on GT, one sees a graph showing its popularity over time. The numbers reflect how many searches have been done for a particular term. [3]

It is reasonable to assume that the GT volume, indicating the level of people's interest, is correlated with stock price. If such correlation exists, stock price could possibly be better predicted with the assistance of GT data. The main focus of this section is to predict the stock price with the assistance of GT data, taking Apple Inc. as an example. We collect the weekly stock price data, which serves as a surrogate for sensor data in our example, and weekly GT data [4] using the search keywords "Apple Inc." + "Apple stock" [5] from Oct. 3 2004 to Jun. 6 2014. The normalized stock price and GT data can be seen in Figure 2.

Most financial studies involve returns, instead of prices, of assets mainly because return series has more attractive statistical properties [15]. Thus, we convert the stock price data and GT data to return series [6], as shown in Figure 3, for the following analysis to be conducted.

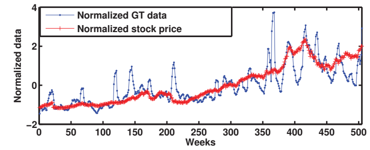

Fig. 2. Normalized weekly stock price and google search volume of Apple Inc. (from Oct. 3 2004 to Jun. 6 2014).

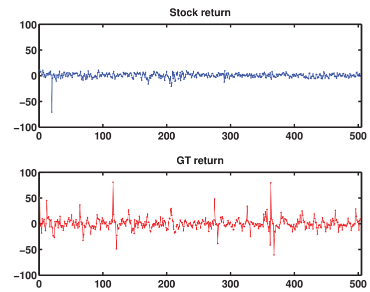

Fig. 3. Stock returns and GT returns (multiplied by 100).

Let the return series of stock be denoted as $\{x_{1,l}\}_{l=1}^{L}$ and the return series of GT be denoted as $\{x_{2,l}\}_{l=1}^{L}$, and in this example $L = 505$. We aim at predicting the stock returns based on its historical data with the assistance of the GT returns i.e., $\hat{x}_{1,l+1} = g(\mathbf{x}_1^{(l)}, \mathbf{x}_2^{(l)})$, by first constructing the conditional joint distribution of $f(x_{1,l+1}, x_{2,l} | \mathbf{x}_1^{(l)}, \mathbf{x}_2^{(l-1)})$. It has to be noted that the two time series are intentionally misaligned by one time instant, and it will be justified later by the non-negligible dependence among the standardized residuals of the misaligned returns.

We first model each of the two returns individually. It can be concluded from Figure 4 that both of the two returns have non-zero AR lag and MA lag, and an ARMA up to the order $(5,5)$ is good enough for modeling the conditional mean part in Eq. (4). Using AIC, an ARMA(5,5) model is selected for stock returns and ARMA(4,5) is selected for GT returns. When BIC is applied, ARMA(5,5) model is chosen for the stock returns and ARMA(1,5) model is selected for GT return.

We conduct Engle's ARCH test [7] for the existence of time-varying conditional variance on the residual series $\{z_{n,1}, \ldots, z_{n,L}\}$. For the stock returns, the null hypothesis is accepted indicating that the variance of the residual series is not time-varying. The result is also supported by Figure 5, in which the stock returns do not have either GARCH term (M) or ARCH term (N), namely constant variance model can be applied to the residuals of stock returns. As to the GT returns, GARCH models up to order $(5,5)$ are considered, and by AIC,

[2] The definition of EDF is scaled by $1/(L+1)$ rather than $1/L$. This has no effect asymptotically, and it is useful in keeping the estimated CDF away from the boundaries of the unit interval, where some copula models diverge.

[3] The numbers don't represent absolute search volume numbers, because the data is normalized and presented on a scale from 0-100. Each point on the graph is divided by the highest point and multiplied by 100.

[4] Data Source: Google Trends (www.google.com/trends).

[5] "+" represents that the results can include searches containing the words Apple Inc. OR Apple stock.

[6] Let $\{y_l\}_{l=1}^{L}$ denote the original time series, the returns in this paper are calculated as follows: $\nabla y_l = y_l - y_{l-1}$

[7] Engle's ARCH test is a Lagrange multiplier test to assess the significance of ARCH effects [13].

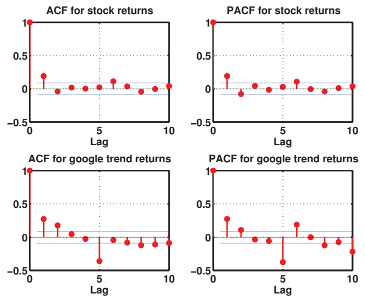

Fig. 4. ACF and PACF of two returns

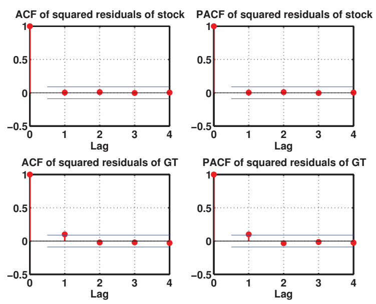

Fig. 5. ACF and PACF of the squared residuals of the two returns

a GARCH(1,1) is selected, while a GARCH(0,1) is selected by BIC.

After the marginal models are selected and corresponding parameters are estimated, the estimated standardized residuals $\{\hat{\epsilon}_{1,1}, \ldots, \hat{\epsilon}_{1,L}\}$ and $\{\hat{\epsilon}_{2,1}, \ldots, \hat{\epsilon}_{2,L}\}$ are obtained through Eq. (8). By checking Kendall's τ [8] for different alignments of the two data sets, i.e., $\{\hat{\epsilon}_{1,1+d}, \ldots, \hat{\epsilon}_{1,L}\}$ and $\{\hat{\epsilon}_{2,1}, \ldots, \hat{\epsilon}_{2,L-d}\}$, we find that $d = 1$ renders the largest $\tau = 0.047$. Thus, the strongest dependence exists between $\epsilon_{1,l+1}$ and $\epsilon_{2,l}$. The dependence between the standardized residuals of stock data at $l+1$ and GT data at l allows us to utilize the current GT data to predict the stock of the next week. And the best copula that is selected to model the dependence between the standardized residuals is Frank copula.

The optimal estimator that minimizes the mean square error (MSE) is the conditional expectation as follows

$$\begin{aligned} \hat{x}_{1,l+1} &= \mathbb{E}[X_{1,l+1} | \mathbf{x}_1^l, \mathbf{x}_2^l] \\ &= \mu_1(\mathbf{x}_1^l | \hat{\theta}_1) \\ &+ \sigma_1 \int f_{\epsilon_1}(\epsilon_1) c^*(F_{\epsilon_1}(\epsilon_1), F_{\epsilon_2}(\epsilon_{2,l}) | \hat{\boldsymbol{\theta}}_d) d\epsilon_1 \end{aligned} \tag{13}$$

where σ_1 is not time-variant since we have shown that there is no conditional heteroskedasticity in the stock returns.

In each of the 30 trials that have been conducted, we use the first l data points, $\{x_{1,2}, \ldots, x_{1,l}\}$ and $\{x_{2,1}, \ldots, x_{2,l-1}\}$ as the training set for the estimation of the model parameters

[8] Kendall's τ is a non-parametric rank-based measure of dependence ranging from -1 to 1.

TABLE II
MSE OF DIFFERENT PREDICTION APPROACHES

Approach	AIC & Frank Copula	BIC & Frank Copula	$c* = 1$
MSE	6.1800	6.2348	6.5960

$\{\boldsymbol{\theta}, \boldsymbol{\phi}\}$, and use the next data point $x_{1,l+1}, x_{2,l}$ as testing data to check the square error of our estimator. The MSE is calculated as follows

$$MSE = \frac{1}{30} \sum_{l=476}^{505} (x_{1,l} - \hat{x}_{1,l})^2 \tag{14}$$

Compared with the approach that assumes intermodal independence between the two returns, i.e., setting $c^* = 1$ in Eq. (13), we are able to reduce the MSE by around 5% as shown in Table II by capturing the intermodal dependence and exploiting the dependence in the process of prediction. Due to encouraging results, further work along these lines is underway.

REFERENCES

[1] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: Real-time event detection by social sensors," in *Proceedings of the 19th International Conference on World Wide Web*. New York, NY, USA: ACM, 2010, pp. 851–860.

[2] S. Asur and B. A. Huberman, "Predicting the future with social media," in *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on*, vol. 1. IEEE, 2010, pp. 492–499.

[3] D. Wang, L. Kaplan, H. Le, and T. Abdelzaher, "On truth discovery in social sensing: A maximum likelihood estimation approach," in *Proceedings of the 11th International Conference on Information Processing in Sensor Networks*, ser. IPSN '12. New York, NY, USA: ACM, 2012, pp. 233–244.

[4] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe, "Predicting elections with twitter: What 140 characters reveal about political sentiment." *ICWSM*, vol. 10, pp. 178–185, 2010.

[5] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011.

[6] X. Wang, M. S. Gerber, and D. E. Brown, "Automatic crime prediction using events extracted from twitter posts," in *Social Computing, Behavioral-Cultural Modeling and Prediction*. Springer, 2012, pp. 231–238.

[7] S. G. Iyengar, "Decision Making with Heterogeneous Sensors - A Copula Based Approach," Ph.D. dissertation, Syracuse University, Syracuse, NY, August 2011.

[8] H. He and P. Varshney, "Fusing censored dependent data for distributed detection," *Signal Processing, IEEE Transactions on*, vol. 63, no. 16, pp. 4385–4395, Aug 2015.

[9] B. Schweizer and A. Sklar, *Probabilistic Metric Spaces*. New York: North Holland, 1983.

[10] R. Nelsen, *An Introduction to Copulas*, 2nd ed. New York: Springer, 2006.

[11] A. Patton, "Copula methods for forecasting multivariate time series," *Handbook of economic forecasting*, vol. 2, pp. 899–960, 2012.

[12] D. C. Montgomery, C. L. Jennings, and M. Kulahci, *Introduction to time series analysis and forecasting*. John Wiley & Sons, 2011, vol. 526.

[13] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation," *Econometrica: Journal of the Econometric Society*, pp. 987–1007, 1982.

[14] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of econometrics*, vol. 31, no. 3, pp. 307–327, 1986.

[15] R. S. Tsay, *Analysis of financial time series*. John Wiley & Sons, 2005, vol. 543.