Understanding Building Vibration with Event Detection and Directed Information based Causal Analysis

Yixin Zheng[1], Wenzhu Zhang[2], Jiantao Jiao[3], Lin Zhang[1]

Abstract— **Recording and analysing building vibration has been proved to be an effective approach to Structural Health Monitoring. Based on a vibration monitoring system, this paper focuses on analyzing building vibration through detection of representative events. We propose an event detection method combining signal processing and machine learning techniques, which is demonstrated to be helpful to extract and understand the vibration signals caused by significant environmental events such as nearby train passings. Furthermore, we apply the recently-developed directed information estimators in Jiao, 2012 to study mechanic wave propagation patterns inside the building when representative events happen through causal analysis. The results demonstrate clear patterns of building's vibration under different conditions, and provide insights for monitoring, assessing and securing the structure's health.**

I. Introduction

In recent years, there has been an increasing interest in the adoption of sensor network techniques for Structural Health Monitoring (SHM), which enables the detection of dynamic structural changes and therefore secures the building usage. As one of the major approaches in this field, accelerometer-based vibration monitoring has attracted much attention, and a number of practical systems have been built to obtain structural vibration data. For example, wired sensor networks have been deployed on tall buildings [1] and suspension brigdes [2] for vibration monitoring. Furthermore, solutions using wireless sensor networks for SHM have also matured, making it possible to reliably collect high precision vibration data at a high rate with low costs [3].

While the studies of vibration data acquisition systems have significantly progressed, our tools to analyze and understand the vibration data are still far from powerful. Applying straight-forward frequency analysis on the vibration data and comparing its results with model parameters is a common practice [1]–[4], which does help verify the system but is not sufficient for us to understand how and why the building's vibration characteristics change [5]. When buildings experience abnormal conditions such as harsh loadings, seismic events or severe environmental events, more tools are needed to identify the events, analyze their influence on the building, and finally determine the building's health condition in preparation for further actions. After all, such significant events are most crucial to building's safety and should be understood more deeply.

Recent advances in information theory concerning directed information provides a new perspective for vibration data analysis. Directed information [6] is a measure that quantifies the causal relationship between different stochastic processes and it plays an important role in such fields as information processing, portfolio theory and biology [7]. While its theoretical properties have been extensively studied, it's not until recently that practical estimators of directed information were developed and proved to be helpful in analyzing various data sets [8], [9], which makes it possible to carry out causal analysis on vibration data based on estimated directed information.

In this paper, we propose a scheme to detect significant environmental events and analyze their influence on the building's vibration characteristics based on directed information estimation. Specifically, we make the following contributions:

- We deploy a vibration monitoring system in an 11-floor office building which is located near a railway and is therefore observably influenced by train passings. With this vibration data acquisition system in its unique location, we build a solid basis for further data analysis and provide an example showing how significant environmental events can affect building vibration characteristics.
- We propose an event detection method for vibration monitoring systems to identify the influence of relevant events. We utilize signal processing techniques such as transform analysis to preprocess the vibration data and extract its features, after which the supervised learning methods are applied to classify patterns representing different events based on prior training. We evaluate this method with our data set and demonstrate that precise detection can be achieved with proper algorithms.
- We apply directed information estimators to study the building's vibration characteristics under the influence of train passings. With its ability to quantify the causal relationship between stochastic processes, we characterize the influence among sensors at different locations to study the propagation of waves stimulated by trains. Results reveal obvious patterns that provide insights to help monitor, assess and secure the structure's health.

The remainder of the of paper is organized as follows. Section II Introduces the fundamentals of directed information and its estimation. Section III explains the system

[1] Yixin Zheng and Lin Zhang are with Department of Electronic Engineering, Tsinghua University, Beijing, P.R.C zyxdaydreams@gmail.com, linzhang@tsinghua.edu.cn

[2] Wenzhu Zhang is with State Information Center of China, Beijing, P.R.C zhangwenzhu@mx.cei.gov.cn

[3] Jiantao Jiao is with Department of Electrical Engineering, Stanford University, Stanford, CA, 94305, USA jiantao@stanford.edu

deployment. Section IV focuses on detection method and its evaluation, while section V concentrates on applying causal analysis on extracted events. Section VI Concludes the paper.

II. Directed Information based Causal Analysis

Directed information was first introduced by Marko [6] and Massey [10] to characterize feedback channels. As a measure to quantify causal influence between two processes, it serves as an alternative to widely-used Granger causality test [11], with a significant improvement that unlike Granger causality test, directed information approach does not require the process pair to be jointly Gauss-Markov [12]. As a result, its use was later extended beyond information theory to biology [13], portfolio theory, data compression and hypothesis testing [7].

In this paper, we introduce directed information to assist vibration data analysis. By measuring the directed information between each pair of processes (that is, the output series of a vibration sensor), we obtain the likelihood of influence between them in both directions, which are direct indicators of their causal relationship.

Before we carry on to analyze the vibration data, a brief introduction to the fundamentals and practical estimation algorithms of directed information is presented in the following subsections. Note that the introduction is by no means comprehensive, and please refer to literature [9], [14] for details.

We use following notations in this section. Uppercase letters $X,Y,...$ refer to random variables , while lowercase letters $x,y,...$ refer to their values. Taking n as the length of the source data series, X^n denotes the n-tuple random variables $(X_1,X_2,...,X_n)$ and x^n denotes $(x_1,x_2,...,x_n)$. The alphabets of X is in its calligraphic form, $\mathscr{X}$, whose cardinality is denoted as $|\mathscr{X}|$.

A. Directed Information

The definition of the *directed information* from X^n to Y^n is as follows:

$$\begin{aligned} I(X^n \to Y^n) &\triangleq \sum_{i=1}^{n} I(X^i;Y^i \mid Y^{i-1}) \\ &= H(Y^n) - H(Y^n \parallel X^n), \end{aligned} \tag{1}$$

where $H(Y^n \parallel X^n)$ denotes the *causally conditional entropy* [14]:

$$H(Y^n \parallel X^n) \triangleq \sum_{i=1}^{n} H(Y_i|Y^{i-1},X^i) \tag{2}$$

Note that unlike *mutual information*, which is a well-known concept in information theory, directed information is not symmetric, e.g. generally $I(X^n \to Y^n) \neq I(Y^n \to X^n)$.

Furthermore, the *directed information rate* [14] is defined as

$$\bar{I}(X^n \to Y^n) \triangleq \lim_{n\to+\infty} \frac{1}{n} I(X^n \to Y^n) \tag{3}$$

For the proof to the limit's existence and its relationship with entropy rate and causally conditional entropy rate, please refer to [9].

B. Directed Information Estimation Algorithms

In spite of its thoroughly studied theoretical properties, it's also of great importance to develop practical estimation algorithms for directed information. The problem is to estimate the directed information rate between two processes, $\bar{I}(X^n \to Y^n)$, based on source data series whose length are usually limited.

As a key in achieving this goal, the Context Tree Weighting (CTW) algorithm [15], a widely adopted sequential probability assignment, offers an effective and reliable method to derive universal probability assignment from data with unknown model and unknown parameters, therefore enabling the estimation of various information measures, including directed information. Furthermore, CTW also enjoys advantages that its computational and storage complexity are both linear in the data length, and the algorithm provides the probability assignment directly. The estimators we apply in this paper, which are introduced and evaluated in Literature [9], are all based on CTW.

The applied 4 directed information estimators as follows:

$$\hat{I}_1(X^n \to Y^n) \triangleq \hat{H}_1(Y^n) - \hat{H}_1(Y^n \parallel X^n) \tag{4}$$

$$\hat{I}_2(X^n \to Y^n) \triangleq \hat{H}_2(Y^n) - \hat{H}_2(Y^n \parallel X^n) \tag{5}$$

$$\begin{aligned} &\hat{I}_3(X^n \to Y^n) \triangleq \\ &\quad \frac{1}{n}\sum_{i=1}^{n} D(Q(y_i|X^i,Y^{i-1}) \parallel Q(y_i|Y^{i-1})) \end{aligned} \tag{6}$$

$$\begin{aligned} \hat{I}_4(X^n \to Y^n) \triangleq &\frac{1}{n}\sum_{i=1}^{n} D(Q(x_{i+1},y_{i+1}|X^i,Y^i) \\ &\parallel Q(y_{i+1}|Y^i)Q(x_{i+1}|X^i,Y^i)) \end{aligned} \tag{7}$$

where

$$\hat{H}_1(Y^n \parallel X^n) \triangleq -\frac{1}{n}\log_2 Q(Y^n \parallel X^n) \tag{8}$$

$$\hat{H}_2(Y^n \parallel X^n) \triangleq \frac{1}{n}\sum_{i=1}^{n} f(Q(x_{i+1},y_{i+1}|X^i,Y^i)) \tag{9}$$

and Q denotes the probability assignment derived with CTW method based on source data, which consists of a set of conditional probability mass function $Q(x_i|x^{i-1})$ for every $x^{i-1} \in \mathscr{X}^{i-1}$. Note that $Q(x_i|X^{i-1})$ denotes $Q(x_i|x^{i-1})$ evaluated for the random sequence X^{i-1}, and such notations as $Q(y_i|X^i,Y^{i-1})$ or $Q(Y^n \parallel X^n)$ are used accordingly. Here function f is defined as mapping from a joint probability mass function $P(x,y)$ to the corresponding conditional entropy:

$$f(P) \triangleq -\sum_{x,y} P(x,y)\log_2 P(y|x) \tag{10}$$

TABLE I
PROPERTY OF DIRECTED INFORMATION ESTIMATORS

	Support	Rates of Convergence
$\hat{I}_1$	$(-\infty, +\infty)$	$O(n^{-1/2}\log_2 n)$
$\hat{I}_2$	$[-\log_2 \lvert\mathscr{Y}\rvert, \log_2 \lvert\mathscr{Y}\rvert]$	$O(n^{-1/2}(\log_2 n))^{3/2}$
$\hat{I}_3$	$[0, +\infty)$	–
$\hat{I}_4$	$[0, +\infty)$	–

Literature [9] proves the convergence of the four estimators to the directed information rate when length of the data series n tends to infinity, mainly in almost sure and L_1 sense. Furthermore, for estimators $\hat{I}_1$ and $\hat{I}_2$, the near-optimal rates of their convergence in the minimax sense are also given, under some mild conditions.

Apart from convergence property, the bounds of the estimators are also of great significance, since the directed information rate $\bar{I}(X^n \to Y^n)$ is known to be nonnegative and upper-bounded by $\log_2 |\mathscr{Y}|$. According to literature [9], while estimator $\hat{I}_1$ suffers the disadvantage of having no upper bound, $\hat{I}_2$ avoids it with an approach to 'smooth' the entropy estimate, yielding it with an upper bound of $\log_2 |\mathscr{Y}|$. However, both $\hat{I}_1$ and $\hat{I}_2$ have an undesirable nonzero probability of being negative. To overcome this shortcoming, $\hat{I}_3$ and $\hat{I}_4$ are proposed, taking forms that are always nonnegative.

A summary of properties of the four directed information estimators is given in TABLE I.

III. SYSTEM SETUP

We deploy a vibration sensor network in an 11-floor office building to study its vibration characteristics under various environmental conditions.

In order to record and analyze the building's vibration under significant environmental events, the location of the building is chosen so that it neighbors a railway with only a short distance of $80-140m$ in between, as shown in Fig.1. More than ten trains with heavy loadings including cargos and hundreds of passengers pass by the building daily, casting influence that are intuitively observable by instruments, yet not strong enough to be directly perceived by human.

A network of accelerometers spread out in the building to record its vibration at different locations and directions. The precision of a single-axis sensor is $10^{-5}m/s^2$, with a highest sample rate of 1024 samples per second. Accelerometers are divided into 16 groups and deployed on the 1^{st}, 5^{th}, 8^{th} and 11^{th} (top) floor of the building, with 4 groups on each floor, whose horizontal positions are illustrated in the vertical view in Fig.1. On the 1^{st} and 11^{th} floor, each group consists of 3 co-located single-axis sensors to monitor the acceleration of west-east, north-south and vertical direction. On the 5^{th} and 8^{th} floor, however, each group consists of only 2 co-located single-axis sensors to monitor the acceleration of west-east and north-south direction. In total, 40 accelerometers are deployed. Coaxial wires are used to connect sensors to a data

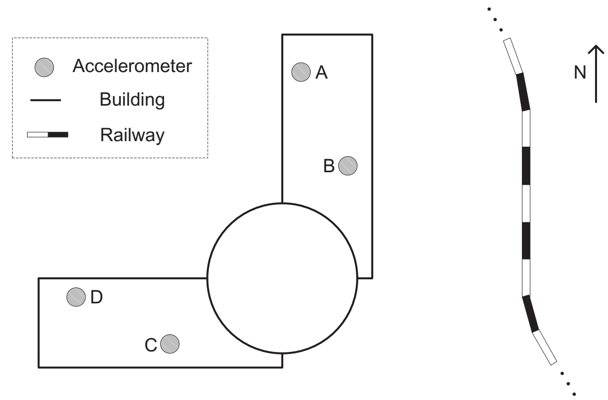


Fig. 1. Building location and sensor deployment (vertical view)

collector at each floor, which collects and sends the vibration data to a central server through local area networks.

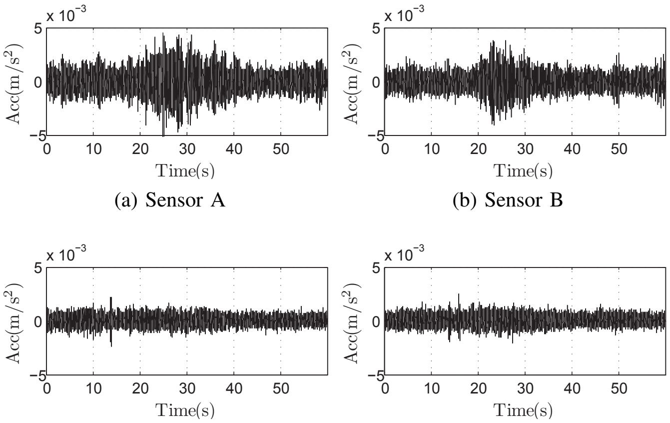
IV. EVENT DETECTION

Our main objective is to study the causal relationship between each pair of vibration data series at different locations, under the influence of significant environment events. Before the directed information estimators are applied on vibration data, a method is needed to identify the such influence from the highly dynamic vibration data.

Event detection based on vibration data is hardly a new idea. One well-studied topic is human activity detection based on wearable accelerometers. Literature [16] first formulates such detection as a classification problem and proposed a machine-learning based approach, with evaluation on vibration data produced by human-worn accelerometers to identify common human activities such as walking, running, standing and sitting. Further studies on event detection with wearable devices such as [17] and [18] provide improved performance. An alternative to wearable accelerometers exists as well, which is usually called 'passive detection', since in such approach the sensors are deployed in the environment instead of on the detection target. For example, [19] focuses on detecting falls of the elderly with a single accelerometer recording the floor's vibration. More generally, vibration data based event detection falls into the larger category of 'temporal data mining', where surveys such as [20] and [21] provide comprehensive overview of problems and solutions.

In spite of the previous research, we meet and address new challenges in our unique scenario.

1) It's much harder to distinguish the vibration stimulated by our detection target from interferences and noise. Unlike previous works, where relevant signals are significantly larger than background readings, the signals caused by environmental events in our scenario are just comparable to the building's natural vibration, if not weaker. To make the matter worse, the vibration signals pass all the way through complex media including ground and the building structure to the sensor, and end up with hardly any observable pattern. Fig. 2



(a) Sensor A (b) Sensor B

(c) Sensor C (d) Sensor D

Fig. 2. The output of 4 vertical accelerometers on the 11^{th} floor. A train passes by at around 20s and its influence lasts for about 10s. The locations of the sensors A to D is illustrated in Fig. 1.

shows a fragment of the signals that lasts for 60s at a sample rate of 200Hz, which includes a part influenced by trains and demonstrates the difficulty to identify such influence in our scenario. This is especially the case for signals shown in Fig. 2c and Fig. 2d, because their locations are further away from the railway (Fig. 1). Compared to the clear and regular pattern studied in previous works, especially in human activity recognition, this is clearly a challenge.

To address this problem, we use transform analysis to extract features that are otherwise opaque in the time domain (see subsection A), a widely-adopted technique in temporal data mining [20]. We further rely on supervised learning techniques to discover patterns from these complex features and finally identify the influence of our detection target (see subsection B).

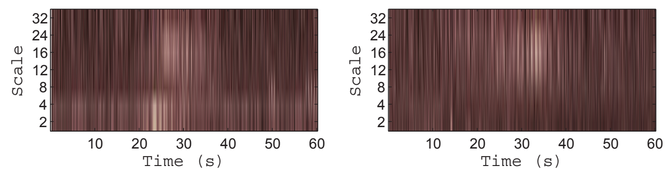
2) The observation of events and their influence on the sensors are not simultaneous, depending on where and how we observe. This happens because the detection targets and the sensors are not co-located, which was not the case in previous works, especially for human activity recognition. For previous works concerning ‘passive detection’, on the other hand, their detection targets are close to the sensors and produce strong, recognizable signals, which makes data labeling an easy task. In our scenario, however, a certain delay generally exists in between and are too significant to ignore, which makes it hard to label the vibration data. This presents a significant challenge to the application of supervised learning technique, which largely depends on labeled data to train classifiers.

We propose a solution to this problem in subsection B that determines the delay between the observed events and their influence on sensors using cross-correlation analysis, therefore enabling reliable training data acquisition.

TABLE II
FOURIER ANALYSIS OF THE VIBRATION DATA IN FIG. 2A

Case		Fourier Analysis Results[a]					
		f_1	PSD_1	f_2	PSD_2	f_3	PSD_3
No Train	I	45	6.57	31	0.79	68	0.50
	II	46	36.79	17	1.08	67	0.52
	III	90	25.35	46	9.28	87	2.83
	IV	46	15.23	11	0.44	91	0.36
Train	I	45	33.25	83	0.93	17	0.64
	II	46	6.83	11	1.99	13	0.59
	III	11	2.50	47	2.11	44	1.63
	IV	46	17.79	69	0.37	16	0.30

[a] $f_i(Hz)$ and $PSD_i(1*10^{-3}m^2/s^3)$ are the frequency and power spectral density of the i^{th} largest peak, respectively.



(a) Sensor B (b) Sensor C

Fig. 3. Wavelet analysis result of the vibration data. (a) and (b) correspond to the analysis result of sensor B (Fig. 2b) and sensor C (Fig. 2c), respectively.

A. Preprocessing and Feature Extraction

Fourier analysis is recognized as the most widely-adopted transform analysis method, which enjoys the advantage of low computation complexity (with FFT) and straight-forward physical interpretation. Here we use fourier analysis to investigate the power spectral density (PSD) of the vibration signal produced by sensor A (Fig. 2a), as shown in TABLE II. We apply a moving window of 1s (that is 200 data points, at the sample rate of 200Hz) on the vibration data, calculate the corresponding PSD with FFT algorithm, and record the frequency and PSD of the 3 largest peaks. We choose some representative cases in the middle of the ’no train’ period as well as the ’train’ period, which we label manually. The results demonstrate that, while the fourier analysis does reveal some differences the train brings, such as larger peaks around 10Hz, it’s still hard to do classification based on any simple rule.

Wavelet analysis [22] provides another transform analysis tool that measures the similarity between a given signal and an analyzing function. Unlike fourier analysis which uses an infinite analyzing function $e^{j\omega t}$, wavelet analysis utilizes the ‘wavelets’ that are localized in both time and frequency domain, extracting the signal’s frequency dynamics as well as time dynamics. This is a significant advantage in our scenario, because the influence of the detection targets changes in both time and frequency aspects. Furthermore, wavelet analysis has low computation complexity and avoids windowing problems.

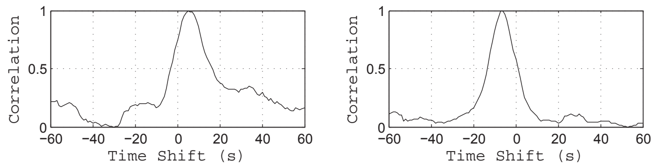


Fig. 4. Cross-correlation between observation sequence and influence sequence

Fig. 3 presents the wavelet analysis results of two vibration data series produced by sensor B and C (Fig. 2b and 2c, respectively) using Continuous Wavelet Transform (CWT), showing scales from 2 to 32. The lighter the color is, the larger the coefficients are, indicating stronger similarity between the signal and the analyzing function. We can clearly observe lighter areas in both figures, ranging roughly from 20s to 40s in time and from 12 to 24 in scale, which reflects the influence of the passing train. Note that for the 'db3' wavelet we use, a scaling factor of 16 stretches it's frequency to 10Hz, which is consistent with our findings from fourier analysis. What's interesting is that even for the signal in Fig. 2c, where hardly anything can be found in time domain, wavelet analysis succeeds in extracting its distinctive features. Meanwhile, the low scale (high frequency) parts for sensor B are also unique and might inspire other findings.

This subsection applies transform analysis techniques on the vibration signals and demonstrates their capability of extracting distinctive features for event detection. In this paper we only investigate two of these techniques and leave other possible choices to future work. After this preprocessing procedure, the machine learning techniques can be applied on these features to realize event detection, as shown in the next subsection.

B. Supervised Learning and Classification

To apply supervised learning techniques on the vibration data, the first step is to obtain labeled data to train the classifiers. For the reasons stated in the previous section, simply observing the events and marking the corresponding signals doesn't work in our scenario because of the delay that generally exists between the observation of events and their influence on the sensors. We therefore propose a method to determine this delay with cross-correlation analysis.

1) Monitor the railway, examine if there are train passings, and obtain a value for each time fragment (e.g. 1s) indicating train's existence in observation. We name it 'observation sequence'. Treating trains heading different directions separately, we get one 'observation sequence' for each direction.
2) Divide the vibration data into fragments of the same length and map the data in each fragment into a single value indicating the likelihood that this fragment is influenced by a train. With the knowledge gained from transform analysis in the previous subsection, we use transform coefficients to determine this mapping (e.g, the wavelet coefficients with scale 16). This is by no means a precise classification, but a fair estimation that is good enough for the purpose of this subsection. We name it 'influence sequence', which has the same length with 'observation sequence'.
3) Calculate the cross-correlation between each 'event sequence' and the 'influence sequence'. The time shift with which the maximum cross-correlation value appears is regarded as the most possible delay between the two sequences.

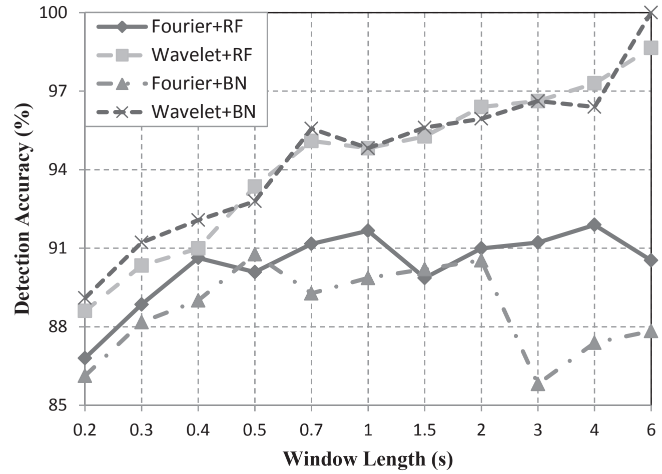


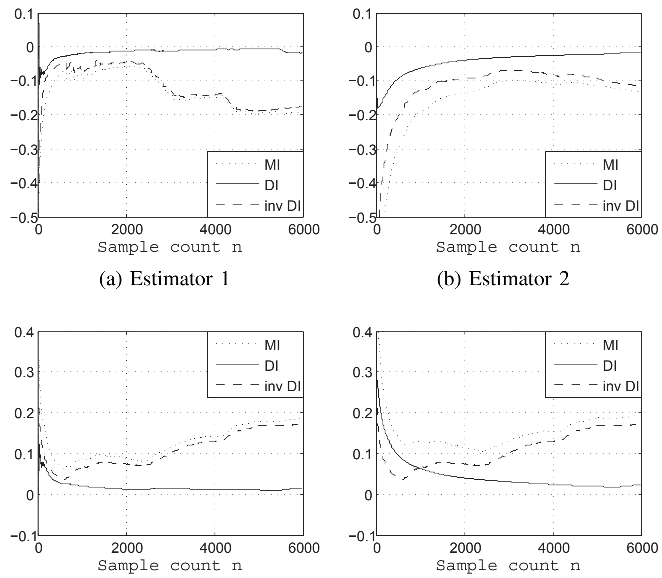
Fig. 5. Detection accuracy evaluation. RF is short for Random Forest algorithm and BN for Bayesian Network algorithm.

The result of cross-correlation analysis is shown in Fig. 4, which demonstrates the significant delay between the observation of events and their influence on the sensors. It turns out that our observation appears 5s earlier for trains heading north, and 6s later for the opposite. Combining this delay with the 'observation sequence' we already have, we solve the data labeling problem.

Next we apply some of the widely-recognized supervised learning algorithms on our labeled data to train classifiers for event detection. Due to the lack of space, in this paper we only consider two of the applicable algorithms, the Random Forest algorithm [23] and the Bayesian Network algorithm [24]. The former is a decision tree style algorithm that uses randomized subsets to train a forest of decision trees and choose the mode of their outputs as the classification result. The latter, on the other hand, learns the conditional dependencies of the random input variables to form a probabilistic graph, and produce outputs accordingly. We use the Weka platform [25] to apply these algorithms on our data sets, and the evaluation results are presented and discussed in the next subsection.

C. Evaluation

We use the 10-fold cross validation method to evaluate the classification accuracy, which is a widely adopted criterion in machine learning that helps reduce the stress of limited labeled data [26].



(a) Estimator 1 (b) Estimator 2

(c) Estimator 3 (d) Estimator 4

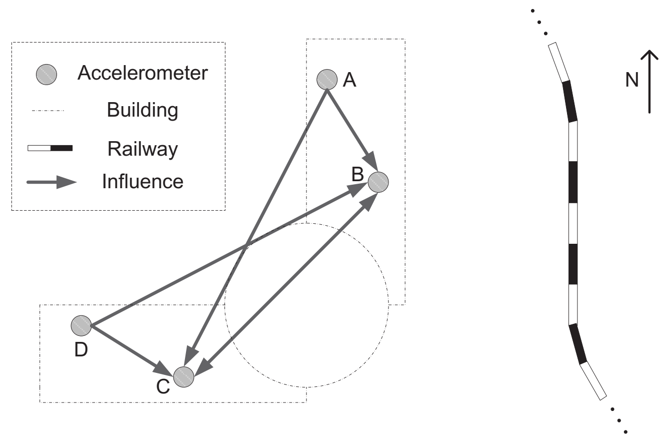
Fig. 6. Directed information with different estimators

Our data set for classifier training and evaluation contains a total of 440s vibration data at the sample rate of 200Hz, including 140s labeled 'Train' by tens of trains and 300s labeled 'No Train', which is produced by the same sensor, e.g. sensor A in Fig. 1, on the 11^{th} floor. We divide the vibration data into fragments (or windows) and extract the feature of each fragment using methods in subsection A. For fourier analysis, the features are the frequency and PSD parameter of the largest 3 peaks, while for wavelet analysis, the features are the average absolute coefficient for each scale level ranging from 1 to 32, with a magnitude of 2.

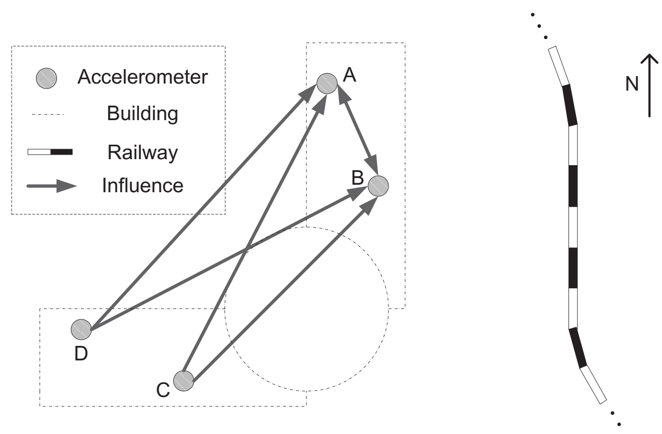
Changing the window length and the classification algorithm, we obtain the detection accuracy evaluation in Fig. 5. Generally, the detection accuracy is above 85% under all setups, and accuracy above 97% is achievable. For wavelet analysis based methods, the accuracy improves as the window length increases, approaching 100% in the end, which was not the case for fourier transform based analysis. The comparison reveals that wavelet analysis performs better, while the Random Forest algorithm and the Bayesian Network algorithm don't make much difference on this data set. Note that here the detection accuracy is measured fragment-wise, not event-wise, because our objective is to determine whether each fragment is influenced by the train and finally form a clear picture of the building's vibration characteristics under environmental event's.

V. Causal Analysis

This section focuses on applying directed information based causal analysis to the vibration data. After we recognize the train's influence on the building's vibration characteristics, we divide the data into two parts depending on whether it's influenced by the train or not. Then the directed information estimators are applied on each part of the data to investigate the causal influences among each pair of the sensors, as well as how they change under significant events' influence.



(a) No Train



(b) Train

Fig. 7. Directed information analysis result

We choose 4 sensors on the 11^{th} floor of the building recording vertical vibration to investigate their influence on each other. Their positions are illustrated in Fig. 1. Since the estimation of directed information requires the alphabet of the data sequence to be finite and relatively small, the sensor readings are discretized before the actual estimation process. Due to computational constraints, the readings are discretized to 5 levels. For more explanation and analysis of discretization before directed information estimation, please refer to literature [9].

Fig. 6 shows the result of four directed information estimators applied on a pair of sensors. In this example, the directed information between sensor B and sensor C is estimated, under the influence of train passings. Each sub-figure corresponds to an estimator introduced in section II, and different estimators turn out to produce obviously different results. Estimator I and II, as stated in section II, produce negative results due to the mismatch between the model and the reality, which is undesirable. On the other hand, the results of estimator III and IV are nonnegative in nature and easier to interpret. In this case, the inverse directed information (inv DI) between sensor B and sensor C is significant, while the directed information (DI) approaches zero, indicating that there exists strong causal influence from sensor C to B when train passes.

Following this example, we carry out directed information estimation between each pair of sensors to obtain a causal graph. Regarding directed information estimation below a threshold (e.g. 0.1) as the influence of noise and ignoring them, Fig. 7 shows the resulting causal graph consisting of causal influence among the vertical sensors with or without train passings. The change of pattern revealed in two causal graphs demonstrates not only the event's influence on the building's vibration characteristics, but also how the building reacts to different kinds of vibration in different ways. Specifically, we make following observations and conjectures:

1) Observation: For natural vibration (no train), the edges of the building influence the center of the building, not otherwise.
   Conjecture: Does natural vibration on the edge cause the center to vibrate?
2) Observation: For forced vibration stimulated by trains, the south wing (perpendicular to the railway) of the building influences the north wing (parallel to the railway), not otherwise.
   Conjecture: Does the south wing cause, or contribute to, the north wing's vibration under train's influence? If so, What's the structural reason for such influence?
3) Observation: Building's vibration demonstrates clear patterns.
   Conjecture: Could we determine the critical district of the building structure under different conditions, therefore help monitor, assess and secure the structure's health condition?

These observations and conjectures need to be further studied and verified with the help and expertise from related fields, concerning structural mechanics and track traffic system's vibration characteristics. Answering these question might provide new perspectives and verify knowledge or models in these fields in return, leading to potential applications to better secure the building structures.

VI. Conclusion

We have made three major contributions focusing on understanding building vibration. We first deploy a vibration monitoring system in an office building near a railway to study the influence of significant environmental events on the building's vibration characteristics. We then propose an event detection method to identify relevant events, with evaluation demonstrating that precise detection can be achieved with proper algorithms. Finally we apply directed information estimators to study the building's vibration characteristics under different conditions, revealing obvious changes in vibration patterns and raising conjectures calling for further research on the causal graphs of vibration and their interpretation.

Many possibilities remain to be studied in future works. First of all, Interpreting causal graphs and relating them with the field knowledge of related domains is an open problem that needs to be addressed. The introduction of directed information based causal analysis casts light on civil engineering and environmental engineering, where traditional model-based analysis could be improved when coupled with techniques in this paper. On the other hand, the event detection algorithms, together with the deployed vibration data acquisition system, still has great potentials to be explored. Other than train passings, local events such as human activities and seismic events such as earthquakes are also possible detection targets with potential value and application.

References

[1] M. Celebi, "Seismic instrumentation of buildings (with emphasis on federal buildings)," *United States Geological Survey (USGS) Report No. 0-7460*, vol. 68170, 2002.

[2] T. Chan, L. Yu, H. Tam, Y. Ni, S. Liu, W. Chung, and L. Cheng, "Fiber bragg grating sensors for structural health monitoring of tsing ma bridge: Background and experimental observation," *Engineering structures*, vol. 28, no. 5, pp. 648–659, 2006.

[3] S. Kim, S. Pakzad, D. Culler, J. Demmel, G. Fenves, S. Glaser, and M. Turon, "Health monitoring of civil infrastructures using wireless sensor networks," in *Information Processing in Sensor Networks, 2007. IPSN 2007. 6th International Symposium on*. Ieee, 2007, pp. 254–263.

[4] J. Ko and Y. Ni, "Technology developments in structural health monitoring of large-scale bridges," *Engineering Structures*, vol. 27, no. 12, pp. 1715–1725, 2005.

[5] S. Doebling, C. Farrar, and M. Prime, "A summary review of vibration-based damage identification methods," *Shock and Vibration Digest*, vol. 30, no. 2, pp. 91–105, 1998.

[6] H. Marko, "The Bidirectional Communication Theory–A Generalization of Information Theory," *IEEE Transactions on Communications*, vol. 21, pp. 1345–1351, 1973.

[7] H. H. Permuter, Y.-H. Kim, and T. Weissman, "Interpretations of Directed Information in Portfolio Theory, Data Compression, and Hypothesis Testing," *IEEE Transactions on Information Theory*, vol. 57, pp. 3248–3259, 2011.

[8] C. J. Quinn, T. P. Coleman, N. Kiyavash, and N. G. Hatsopoulos, "Estimating the directed information to infer causal relationships in ensemble neural spike train recordings," *Journal of Computational Neuroscience*, vol. 30, pp. 17–44, 2011.

[9] J. Jiao, H. Permuter, L. Zhao, Y. Kim, and T. Weissman, "Universal estimation of directed information," *Arxiv preprint arXiv:1201.2334*, 2012.

[10] J. Massey, "Causality, feedback and directed information," in *Proc. Int. Symp. Inf. Theory Applic.(ISITA-90)*, 1990, pp. 303–305.

[11] C. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, pp. 424–438, 1969.

[12] P.-O. Amblard and O. J. J. Michel, "Relating granger causality to directed information theory for networks of stochastic processes," *CoRR*, vol. abs/0911.2873, 2009.

[13] A. Rao, A. Hero 3rd, J. Engel *et al.*, "Using directed information to build biologically relevant influence networks," in *Comput Syst Bioinform Conf (S1752-7791)*, vol. 6, 2007, pp. 145–156.

[14] G. Kramer, "Directed information for channels with feedback," Ph.D. dissertation, University of Manitoba, Canada, 1998.

[15] F. M. J. Willems, Y. M. Shtarkov, and T. J. Tjalkens, "The context-tree weighting method: basic properties," *IEEE Transactions on Information Theory*, vol. 41, pp. 653–664, 1995.

[16] N. Ravi, N. Dandekar, P. Mysore, and M. Littman, “Activity recognition from accelerometer data,” in *Proceedings of the National Conference on Artificial Intelligence*, vol. 20, no. 3. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2005, p. 1541.

[17] L. Bao and S. Intille, “Activity recognition from user-annotated acceleration data,” *Pervasive Computing*, pp. 1–17, 2004.

[18] J. Liu, L. Zhong, J. Wickramasuriya, and V. Vasudevan, “uwave: Accelerometer-based personalized gesture recognition and its applications,” *Pervasive and Mobile Computing*, vol. 5, no. 6, pp. 657–675, 2009.

[19] M. Alwan, P. Rajendran, S. Kell, D. Mack, S. Dalal, M. Wolfe, and R. Felder, “A smart and passive floor-vibration based fall detector for elderly,” in *Information and Communication Technologies, 2006. ICTTA’06. 2nd*, vol. 1. IEEE, 2006, pp. 1003–1007.

[20] C. Antunes and A. Oliveira, “Temporal data mining: An overview,” in *KDD Workshop on Temporal Data Mining*. Citeseer, 2001, pp. 1–13.

[21] S. Laxman and P. Sastry, “A survey of temporal data mining,” *Sadhana*, vol. 31, no. 2, pp. 173–198, 2006.

[22] S. Mallat, *A wavelet tour of signal processing*. Academic Pr, 1999.

[23] T. Ho, “The random subspace method for constructing decision forests,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 8, pp. 832–844, 1998.

[24] N. Friedman, D. Geiger, and M. Goldszmidt, “Bayesian network classifiers,” *Machine learning*, vol. 29, no. 2, pp. 131–163, 1997.

[25] G. Holmes, A. Donkin, and I. Witten, “Weka: A machine learning workbench,” in *Intelligent Information Systems, 1994. Proceedings of the 1994 Second Australian and New Zealand Conference on*. Ieee, 1994, pp. 357–361.

[26] I. Witten, E. Frank, and M. Hall, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2011.