A Prognostics Framework for Health Degradation and Air Pollution Concentrations

Bernard Fong
Prognostics and Health Management Center, City University of Hong Kong
Email: bfong@ieee.org

G. Y. Hong
Unitec, New Zealand

***Abstract*— Air pollution is known to cause a wide range of diseases that have lead to countless cases of premature mortality. Although pollution in a massive scale is uncontrollable by individual inhabitant, the negative impact on health as a direct result on indoor air pollution can be reduced by appropriate preventive actions. Smart home technology allows ventilation to be controlled through condition-based monitoring of the air pollutant concentration in the room. This paper describes a prognostics framework that regulates in-room ventilation, the algorithm can be fed into a smart home control system for prevention of a number of respiratory diseases.**

***Index Terms*—Air pollution, disease control, health monitoring**

I. Introduction

Air pollution is not only a global issues, indoor air pollution due to poor ventilation also causes significant morbidity from respiratory diseases [1]. Indoor air pollution is becoming a serious public health issue as air pollution turns homes into one of the worst environment for living that has been created as a result of burning for cooking and heating. Long term low-level exposure to air pollutants has been known to cause a range of both chronic obstructive pulmonary disease (COPD) and acute respiratory infections (ARI) [2], [3]; mainly caused by combustion, chemical toxins, radon gas and biological organisms [4]. To minimize the health risk of home occupants, a series of longitudinal health data coupled with detailed monitoring of personal exposure is necessary in order to accurately estimate the exposure-response relation for air pollutants from various sources.

This paper investigates the use of a prognostics framework for controlling indoor air pollution in the home by providing an overview of the elementary and critical aspects of linking indoor pollution to chronic disease with prognostics methodology involving the study of day-to-day variations in air pollution and health [5]. A framework is necessary to provide necessary statistical information for air pollution control using smart home technology based on that proposed in [6]. Local awareness in smart home applications such as [7], [8] would allow regulation of air flow based on user location.

II. Indoor Air Qaulity

Air pollution data collected by monitoring of areas such as kitchens and living rooms can be used to regulate indoor air quality through prognostics and air flow management of different locations. The four major parameters to monitor include:

A. Fuel combustion

Indoor air pollution from fuel combustion and acute respiratory illness affect many children in regions where home heating is necessary [9]. Unvented heaters using kerosene can also produce acid aerosols [10]. Combustion induced pollutants not only affect cold climate regions, kitchen gas stoves also produce hazardous gases such as carbon monoxide and nitrogen dioxide.

B. Chemical toxins

Exposure to volatile organic compounds (VOCs) at home is one of the major toxic chemical risks [11]. VOCs are emitted from numerous chemical products used at home. A simple cleaning task, using glue or hobby painting can also generate VOCs given that organic substances are very widely used as ingredients in household products. The heavy metal lead is also a source of chemical hazard in many homes [12]. Discharged into the air through airborne dust and lead-based paint deterioration, inhaled lead particles or dust can risk accumulation in the respiratory system.

C. Ambient radioactivity

Sources of radioactivity in the ambient environment includes the radioactive noble gas Radon; itself being a product of the natural radioactive decay chain of uranium which is found in soil and rocks around the world [13]. Both radon and uranium emit gamma rays and will retain their presence at the same concentrations for centuries to come [14]. Along with other gases such as oxygen and carbon monoxide, radon readily dissolves in the blood and circulates throughout the body.

doi:10.4304/jait.3.1.64-68

D. Biological organisms

Biological contaminants such as bacteria, viruses, molds, pollens, and mites; come from many sources. Bacteria and viruses carried by people, pets and cockroaches; pollens from plants and potent allergen from rodents [15]. The control of these contaminants through the home needs to be analyzed through regulation of relative humidity level to minimize any breeding grounds.

III. Bio-syndromic Surveillance and Computational Modeling

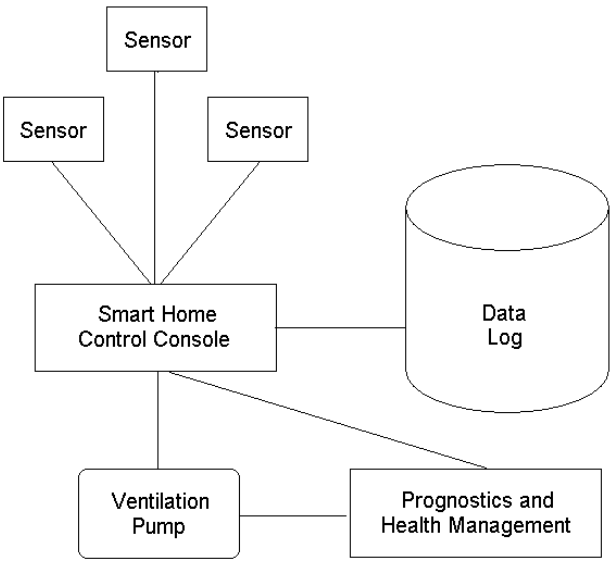

Fig. 1 Smart home control block diagram

An effective yet simple home air pollution surveillance method is needed to provide the necessary information for a smart home system that controls indoor air quality. The block diagram in Fig. 1 shows how indoor air quality can be adaptively regulated by integrating into a smart home control system. Most bio-syndromic surveillance methods can be classified into three regression classes, namely linear regression, Poisson regression and regression with ARIMA (Autoregressive Integrated Moving Average) error structures. Covariates such as day-of-week indicators, seasonal trends with harmonic terms, and holiday indicators are often included in the regression models. Syndromic surveillance may include many evolving data streams with new data sources and the four syndrome categories outlined in Section II to be modeled simultaneously. The Holt-Winters exponential smoothing method is reported in [16] as robust in the sense that reliable, short-term forecasts can be obtained for a broad range of time series. [17] compared the performance of the multiplicative Holt-Winters procedure to adaptive and non-adaptive regression methods which reported a serious limitation of the Holt-Winters method is its inability to easily incorporate external information. Multiple seasonality and over-dispersed count data also impose critical challenges to the exponential smoothing method. In order to model discrete count data in health surveillance, a Dynamic Generalized Linear Model (DGLM) that incorporates covariates such as sub-regions, days of the week, holidays, and seasonality is necessary to take into consideration various attributes of the home, such as the presence or absence of each occupant. The DGLM accommodates hierarchical modeling of spatial data so that spatiotemporal data can be appropriately handled in a systematic framework [18]. When modeling the relationships among heterogeneous health-related data streams, sampling frequencies are often different while frequently sampled streams may have leading power to predict the key performance indicators. The DGLM allows accurate disaggregation of space-time data so that different levels of disease-related information can be integrated and synchronized in combination with the Bayesian network model [19]. This work uses the DGLM to include dynamic covariates for identification of dynamic clusters and their relationships during a disease. A semantic web approach such as that in [20] can be used to identify clusters based on dynamic covariates.

IV. Smart Home Control

Using a Bayesian framework, an algorithm to calibrate and update model parameters can be developed such that the predicted values will be compared to observed values. Any significant differences will be indicated in the application for the smart home control system. Suppose $X = (X_1, X_2, ..., X_n)$ be a sample of size n from the data set of some probability function $F(X)$. The empirical distribution function $F_n(X)$ can be obtained through the available sample $X = (X_1, X_2, ..., X_n)$ by assigning a probability of $1/n$ at each point, $X_1, X_2, ..., X_n$, of the sample. For any given monitored parameter $\theta = \theta(X)$, prediction of the standard error $SE(\hat{\theta})$ can be computed from the standard deviation of the N replications:

$$SE(\hat{\theta}) = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}[\hat{\theta}^{*(j)} - \theta(\cdot)]^2} \quad (1)$$

The air pollution concentration of a given parameter X is $\lambda_X T_0$. Then the hypothesis test can be defined as:

$$H_0 : \mu_X = \lambda_X T_o,\ H_1 : \mu_X > \lambda_X T_o \quad (2)$$

$$\text{Accept } H_0, \text{if } \hat{\mu}_X \le \lambda_X T_0 + \frac{\hat{\sigma}_X}{\sqrt{n}} Z_{1-\alpha}$$

$$\text{Reject } H_0, \text{ if } \hat{\mu}_X > \lambda_X T_0 + \frac{\hat{\sigma}_X}{\sqrt{n}} Z_{1-\alpha}$$

Where n is the sample size, $\hat{\mu}_X$ is the estimate of the mean of the variable X, $\hat{\sigma}_X$ is the estimate of the standard deviation of the variable X, α is the producer's risk, and $Z_{1-\alpha}$ is the ($1-\alpha$) lower percentile point of the Standard Normal Distribution $N(0,1)$.

By applying fuzzy ontology for assessment [21], ventilation control can be activated according to:

Accept H_0 if $H_T \geq 1-\theta$;

Continue monitoring if $\theta < H_T < 1-\theta$;

Reject H_0 and suspend ventilation if $H_T \leq \theta$.

By using prospective estimation for Pollutant Spread Modeling, known information is used to estimate the pollutant spread and the change concentration over time. The following steps will be taken:

A. Hazard identification

The health effects induced by the target pollution source are identified. For example, breathing difficulty due to a given condition in association with an asthma patient.

B. Concentration-response assessment

The relationship between the inhaled amount of pollutants and the probability of accumulation that consequently leads to a disease is mathematically described. For example, the prolonged inhalation of radon gas in the basement may increase the risk of lung cancer while radon decays into a radioactive heavy metal prior to exhalation.

C. Exposure assessment

The exposure levels of the pollutant under a particular scenario are assessed. For example, the use of gas stove for cooking with inadequate ventilation may cause excessive accumulation of carbon monoxide.

D. Infection probability characterization

The probability of being infected for a user is computed based on the exposure levels and the concentration-response relationship. For example, the built-up of pollen may require closing the window and ventilation may worsen the problem.

In an example of assessing the respiratory infection probability due to photochemical pollutants such as nitrogen dioxide can be estimated using different methodologies such as exposure assessment via inhalation pathway. The inhalation exposure (*E*) assessment model proposed in [22] can be used in fluid dynamics modeling for different locations to estimate the risk of infection at different locations and time of pollution exposure. Data collected from various locations by methods such as biological air sampling, can be used to construct a computational model. Computational fluid dynamics will be employed to model the spatial and temporal distribution of the pollutant in the different locations of the home. The Euler-Lagrange multiphase model is used to simulate the airflow and droplet transport [23]. Turbulence closure of the carrier phase can be analyzed by the Reynolds Averaged Navier-Stokes (RANS) approach, using the renormalization group (RNG) *k-ε* turbulence model. Polydispersed discrete particles following the physical size range of cough-generated expiratory droplets is then fed into the model for statistical analysis. This can then be used to regulate the air flow to ensure dilution of pollutants with the mechanism shown in Fig. 2 such that the air pump that regulates air flow for a given room is driven by the analyzed air quality.

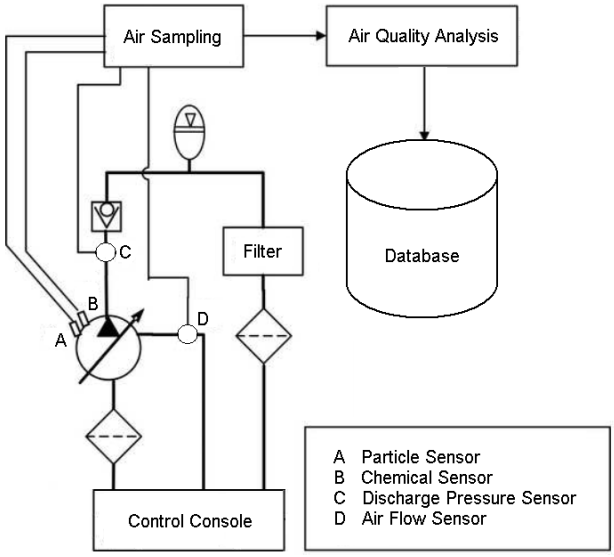

Fig. 2 Ventilation control system

Computational fluid dynamics modeling can be used for predicting the number of particles accumulated in the body that is not exhaled. This is particularly useful for heavy particles and radon that risks radioactive decay into harmful heavy metal substances. Gaussian process models can also be derived for estimating the number of additional pollutants. A Bayesian averaging approach will be used for predicting the number of pathogens in a given room where a fluid dynamics model is not available.

V. A Prognostics Approach for Air Pollution and Health Degradation

To address the issue of indoor air pollution induced health degradation, spatial and spatio-temporal models used to investigate pollutant distribution, pollutant prediction at unmeasured sites, and pollutant prediction over time. One major challenge is to achieve a comprehensive view of air quality by combining time series of multiple pollutants at different rooms, spatio-temporal measurements with different instruments, and different level summary statistics. To integrate multiple data streams and capitalize the complex dependence across time and space, it is of both theoretical and practical importance to develop a hierarchical model with

conditional sub-models defined hierarchically at different levels. The uncertainty is apportioned to different levels and propagated through the hierarchy, which also provides a formal way to borrow strength between various components and improve the precision of statistical inference. This complex parameter structure poses a challenge for statistical inference. In this research we will develop a Bayesian approach to tackle this problem. Markov Chain Monte Carlo (MCMC) algorithms can be used to compute the model by taking advantage of the conditional independence structure in the hierarchical model.

Ventilation controls the relationship between discharge flow and discharge pressure of the air regulated. Fig. 3 shows the simulated flow rate of the regulated air pump at a maximum value of 6 L/min with leakage when the discharge pressure is less than the spring pressure of 1 000 KPa. Fig. 4 shows pump failures distribution due to insufficient inlet pressure.

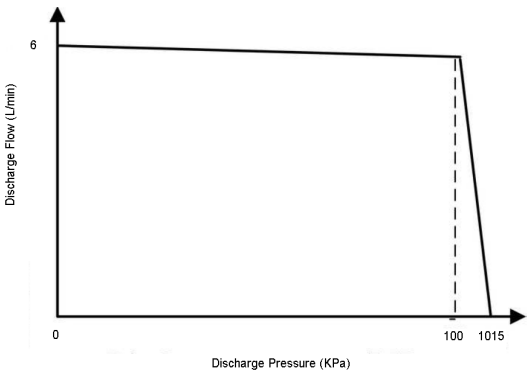

Fig. 3 Ventilation pump discharge flow vs. discharge pressure

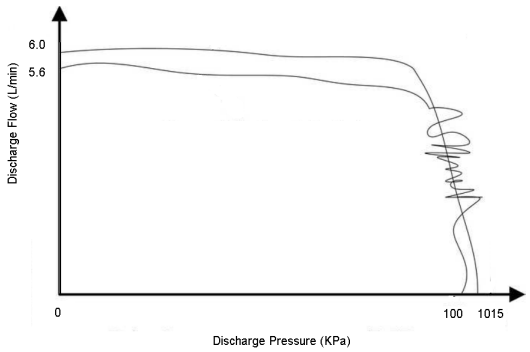

Fig. 4 Ventilation pump under insufficient inlet pressure

VI. Concluding Remarks and Future Work

A control algorithm is described for implementation in a home environment for air quality control through condition-based ventilation. Although the cost of setting up necessary sensors makes it impractical for an average home, it provides a framework for patients with special medical conditions. Appropriate sensors for a particular type of pollutant can be temporarily installed in individual rooms and connected to an existing smart home control system using the method described in [24]. Future improvement can be made with more data for calibration and operational data for reliability assurance with prognostics and health management methodology.

REFERENCES

[1] C Corvalán, S Hales, AJ McMichael, Ecosystems and human well-being: health synthesis, (2005); World Health Organization,.

[2] J Sunyer, et. Al., Air pollution and emergency room admissions for chronic obstructive pulmonary disease: a 5-year study, American Journal of Epidemiology, (1993); 137(7):701-705.

[3] HR Anderson, et. Al., Air pollution and daily admissions for chronic obstructive pulmonary disease in 6 European cities: results from the APHEA project, Eurorean Respiratory Journal, (1997); 10(5):1064-1071.

[4] BH Chen, CJ Hong, MR Pandey and KR Smith, Indoor air pollution in developing countries, World Health Statistics Quarterly (1990); 43(3):127-138.

[5] B Brunekreef and ST Holgate, Air pollution and health, The Lancet (2002); 360(9341):1233-1242.

[6] F Zuo and PHN de With, Real-time embedded face recognition for smart home, IEEE Transactions on Consumer Electronics (2005); 51(1):183-190.

[7] S Kee, KN Ha and KC Lee, A pyroelectric infrared sensor-based indoor location-aware system for the smart home, IEEE Transactions on Consumer Electronics (2006); 51(4):1311-1317.

[8] R Iqbal, A James, J Black and W Poreda, Peripheral Display for Multi-User Location Awareness, Journal of Advances in Information Technology (2010); 1(3):116-126.

[9] M Ezzati and DM Kammen, Quantifying the effects of exposure to indoor air pollution from biomass combustion on acute respiratory infections in developing countries, Environmental Health Perspectives (2001); 109(5):481–488.

[10] P Koutrakis, SLK Briggs and BP Leaderer, heaters using kerosene can also produce acid aerosols, Environmental Science and Technology (1992); 26(3):521-527.

[11] J Manuel, A healthy home environment?, Environmental Health Perspectives (1999); 107(7): A352–A357.

[12] RJ Lioy, NCG Freeman and JR Millette, Dust: a metric for use in residential and building exposure assessment and source characterization, Environmental Health Perspectives (2002); 110(10):969–983.

[13] JAS Adams, WM Louder, G Phair, and D Gottfried, The Natural Radiation Environment, US Department of Energy Nuclear Testing Archive, Accession Number: NV0053511, Document Number: 57452.

[14] WL Roper, Toxicological Profile for Radon, Agency for Toxic Substances and Disease Registry, US

Public Health Service in collaboration with the US Environmental Protection Agency.
[15] N Kalogerakis et. Al., Indoor air quality—bioaerosol measurements in domestic and office premises, Journal of Aerosol Science (2005); 36(5):751-761.
[16] S Makridakis et. Al., The accuracy of extrapolation (time series) methods: results of a forecasting competition, Journal of Forecasting (1982); 1:111-153.
[17] HS Burkom, SP Murphy and G Shmueli, Automated time series forecasting for biosurveillance, Statistics in Medicine (2007); 26:4202–4218.
[18] S Banerjee, AE Gelfand and BP Carlin, Hierarchical Modeling and Analysis for Spatial Data, Chapman and Hall/CRC Press (2004).
[19] F Moauro, and G Savio, Temporal disaggregation using multivariate structural time series models, Econometrics Journal (2005); 8:214-234.
[20] QT Tho, SC Hui and ACM Fong, Semantic web support for customer service, Advances in Electronic Business, Vol. II, Ed. CT Du, E Li, Idea Group, Inc. (2007); 93-115.
[21] QT Tho, SC Hui and ACM Fong, Automatic fuzzy ontology generation for semantic Web, IEEE Transactions on Knowledge and Data Engineering (2006); 18(6):842-856.
[22] GNS To et. Al., A methodology for estimating airborne virus exposures in indoor environments using the spatial distribution of expiratory aerosols and virus viability characteristics, Indoor Air (2008); 18:425-438.
[23] MP Wan et. Al., Dispersion of Expiratory Aerosols in a General Hospital Ward with Ceiling Mixing Type Mechanical Ventilation System. Aerosol Science and Technology (2007); 41:244-258.
[24] N Ansari, C Zhang, R Rojas-Cessa, P Sakarindr and ESH Hou, Networking for critical conditions, IEEE Wireless Communications (2008); 15(2):73-81.