

An integrated service for Personalised Healthcare and Wellness (PHW)

Su-Shing Chen

Systems Biology Lab,
CISE Department,
University of Florida,
Gainesville, FL 32611, USA
E-mail: suchen@cise.ufl.edu

Abstract: At the Systems Biology Lab of the University of Florida, we are developing an integrated system, PHW service, for accessing and storing and monitoring information in personalised medicine. We intend to take a practical approach to build such a system in the framework of clinical and translational sciences of the NIH Roadmap.

Keywords: PHW; personalised healthcare and wellness; personalised medicine; clinical and translational sciences; NIH; genomics; gene regulation; access control; security; privacy; electronic healthcare record; ontology.

Reference to this paper should be made as follows: Chen, S-S. (2010) 'An integrated service for Personalised Healthcare and Wellness (PHW)', *Int. J. Functional Informatics and Personalised Medicine*, Vol. 3, No. 2, pp.144–156.

Biographical notes: Su-Shing Chen received his PhD in Mathematics from the University of Maryland, College Park, in 1970. He has been on the Faculty of several universities as well as Visiting Scientist of some industrial research labs. He started his career at the University of Florida in 1970, and finally retired from there in 2006. He is Emeritus Professor of CISE at UF and CECS at the University of Missouri-Columbia. Still active in research, he is the Director of Systems Biology and affiliated faculty of BMI, Genetics Institute and McKnight Brain Institute of UF. Currently, he runs in Palo Alto, California, an international business and technology consultancy service.

1 Introduction

This paper aims at an integrated approach to personalised medicine that will employ previous research results (references cited), and will innovate the context to 'Clinical and Translational Sciences', which is a key goal of NIH research programs on genomics and proteomics. As well known, personalised medicine is based on various omics (genomics, proteomics, metabolomics, transcriptomics etc.), thus it requires a great deal of data mining and knowledge extraction. The research project includes:

- privacy and security, and access control
- content creation, indexing, information visualisation

- graphical and network models in genomics and proteomics
- wireless devices – iPhones, iPods.

The genetic information will be supplemented and integrated with traditional medical information (e.g., EHR or EMR) and lab measurements as standardised in Logical Observation Identifier Names and Codes (LOINC). The graphs and networks will represent the specific health pathways and gene regulatory networks that are essential to personalised medicine.

The integrated system, PHW service, starts with a novel access control scheme for various stakeholders accessing the data centre services, which monitor and manage users (patients) personal medical information. The information is then input to the users' profiles, which have certain network models representing the genomic and proteomic knowledge. While such knowledge can feedback to users and their physicians or nurses, users may also access themselves. In this case, the information visualisation function facilitates the users' access and utilisation. This paper provides a practical and feasible system for the Clinical and Translational Sciences research. (NIH Roadmap) We will collaborate with the Clinical and Translational Sciences Institute (CTSI) at UF¹ and other institutions for further clinical experiments. This paper establishes the user interface to a PHW data centre, with potential extension to robotics and sensors. The graphical and network models may be integrated to multi-objective optimisation for personalised drug discovery and treatment. This paper employs current off-the-shelf information and communication technologies, providing economic impacts to US companies. The novel access control scheme will be useful to the further development of secure social networks.

2 Privacy, Security and access control

In Chen et al. (2006), we have introduced a novel internet security scheme for access control in real-world applications. It can now be used in this project. In addition to traditional privacy and security issues (e.g., authentication, authorisation and HIPPA²), this scheme provides the additional necessary access control for patients – physicians and nurses – in an appropriate hierarchy. Because potentially millions of users will access billions of objects (documents or records) of information content in complex processes, access control is one of the strategic technologies that will increase the value and utility of the internet and internet-based applications. Traditional security issues deal with users in network domains. In the internet age, our services require further appropriate security mechanisms in information domains (HIPPA). We must address system-level security policy, modelling and enforcement. For example, whether an authenticated user is permitted to access an object, based on its execution domain and membership in some user community? What roles to play? Which objects to access? The PHW services exhibit unique requirements of dependency on user interaction and information sharing. Instead of all users loosely clustered into one category, in such a community of users, additional access constraints imposed on the hierarchy of user roles are needed. The first proposed access control mechanism is called the Role-Based Access Control (RBAC).³ The RBAC mechanism has limitations, because there is no protection of privacy. In our 2003 paper (Chen et al., 2006), we proposed an additional object-based model, or the Boolean Expression Access Control (BEAC) mechanism, to address the information

domain issues in the context of community-based information sharing. Given the PHW domain, the user roles as subjects can better capture the organisational structure of the PHW community of users. Moreover, the object model allows personalisation of objects according to user roles and specific PHW services. Finally, the BEAC mechanism provides flexibility in regulating accesses to objects by roles. Privacy of information can be abused in a healthcare and wellness domain. Even worse, the integrity of the systems could be compromised. Although access control is often described as rules regulating how subjects are allowed to access objects, it can be viewed as information flow control since every access results in flow of information between entities (either or both subject and object). Within the information flow, there are exceptions to the normal information flows. For example, sensitive information content can be allowed to flow from A to B and from B to C, but not permissible to go from A to C directly. This kind of security control is also addressed in our system using the BEAC mechanism (see details in Chen et al., 2006).

Traditionally, security mechanisms include authentication, authorisation and auditing of users in network domains. Authentication (and mutual authentication) is a well-established research area. The entities to be authenticated are typically the identity of subjects. We have included the roles to represent a subject. For example, physicians, nurses and patients are roles. In this proposed system, roles are authenticated in addition to the identity of the subject (e.g., user id and password). The use of roles for authentication of subjects matches the organisational structure more naturally with the information system, and thus can lend a more manageable implementation of the security model. Although a role is equivalent to a group with similar privileges in a conventional system, roles carry semantic meaning that can be used to facilitate efficient and proper assignments of privileges to subjects and consistent conflict resolution when a subject carries multiple and possibly overlapping roles. In general, the number of roles required and to be managed in a system will be significantly smaller than unstructured groups. Authorisation of information access in traditional information systems is typically a static check of what subject (including its group memberships) has the privilege of accessing what objects. The access policies that we consider in this proposal are more dynamic with additional conditions imposed on the access rules. For example, a patient can access certain information only after the physician has examined the medical objects and for limited capacities. In a sense, the access control policy becomes state-dependent. The role-based concept is useful in modelling these types of complex protection policies. Objects to be accessed carry attributes information (both stateful and stateless) and roles are associated with attributes. The association of roles to a subject can change with respect to each particular instance. Furthermore, roles can be hierarchical to represent the organisational structure. For example, it might be more desirable that the role of a hospital counsellor be considered as a subset of the role of a physician. An access control model regulates the access of any active entity (subject) to any passive entity (object) securely. All system attributes are given some security attributes, and a set of access control rules determines how a subject can access an object based on their security attributes. Usually, access control models are divided into mandatory and discretionary access control models. Both are formulated to allow or deny particular access models by subjects to objects. Whereas mandatory models govern reading and writing of information contained in an object by a subject, discretionary models typically allow a richer set of access modes for an object, depending on the type of the object. Both models differ mainly in how access authorisations can be modified. In a mandatory model,

authorisation modifications can only be made by system security administrators through changing the security attributes of subjects or objects, whereas a discretionary model gives a subject some degree of freedom to pass the whole or part of its access privileges for an object to another subject. In real world applications, we have found needs for new models and that both mandatory and discretionary models are not sufficient. For example, there are different exceptions (e.g., transitivity, aggregation, and separation) to the multi-level information flow, which violate the basic properties of lattice of the mandatory access control model.

The BRAC model has been subsequently extended to include states, to remedy the inadequacies of the mandatory and discretionary access control mechanisms. It is a promising method for controlling what information users can make use, programs they can run, and modifications they can make. However, it has some limitations (e.g., overlapping of roles) so that we have to develop another BEAC object-based access control mechanism to complement it. The combination of the two mechanisms is called the role-/object-based access control mechanism in Chen et al. (2006). It is obvious that access control is state dependent of subjects and objects. For example, a decision on a subject accessing an object may depend on previous accesses of some objects by the same subject, or other subjects accessing the same object (a good example is a one-time drug prescription).

With role-based access control, access decisions are based on the roles that individual users have as part of an organisation. Users take on certain assigned roles (such as physicians, nurses, and patients). The process of defining roles depends on how an organisation operates, and requires very careful studies. Access rights are grouped by role name, and the use of resources is restricted to individuals authorised to assume the assigned role. For example, within PHW system the role of physician can include operations to order tests, diagnose diseases, prescribe drugs, and provide therapies; the role of counsellor is limited to advising patients and having information on patients only; the role of nurses only a subset of physicians. The use of roles to control access will be effective for developing and enforcing PHW-specific security policies, and for streamlining the security management process. From the PHW perspective, it is a realistic framework for implementation. In the RBAC mechanism, users are granted membership into roles based on their responsibilities in the service. The operations that a user is permitted to perform are based on the user's role. User membership into roles can be revoked easily and new memberships established as job assignments dictate. Role assignments can be established when new operations are defined, and old operations can be deleted and updated as organisational functions change and evolve. This simplifies the administration and management of responsibilities; roles can be updated in the role file without updating the responsibilities for every user on an individual basis. When a user is assigned to a role, the user can be given no more responsibilities than is necessary to perform the job. The concept of 'least responsibility' requires identifying the user's job functions, determining the minimum set of responsibilities required to perform that function, and restricting the user to a domain with those responsibilities and nothing more. In less precisely controlled systems, this is often difficult or costly to achieve. Some one assigned to a job category may be allowed more responsibilities than needed because it is difficult to tailor access based on various attributes or constraints.

Role hierarchies can be established to provide for the structure of a PHW service. A role hierarchy defines roles that have unique attributes and that may contain other roles; i.e., one role may implicitly include the operations that are associated

with another role. Role hierarchies are a natural way of organising roles to reflect the responsibility: RBAC roles can have overlapping responsibilities, i.e., users belonging to different roles may need to perform common operations. Some general operations may be performed by all community members. In this situation, it would be inefficient and administratively cumbersome to specify repeatedly these general operations for each role that gets created. The role in which the user is gaining membership is not mutually exclusive with another role for which the user already possesses membership. These operations and roles can be subject to organisational policies or constraints. When operations overlap, hierarchies of roles can be established. Instead of instituting costly auditing to monitor access, organisations can put constraints on access through RBAC. This inheritance will be implemented by an object-oriented role database (e.g., the user profiles & roles database).

3 Content creation, indexing, information visualisation

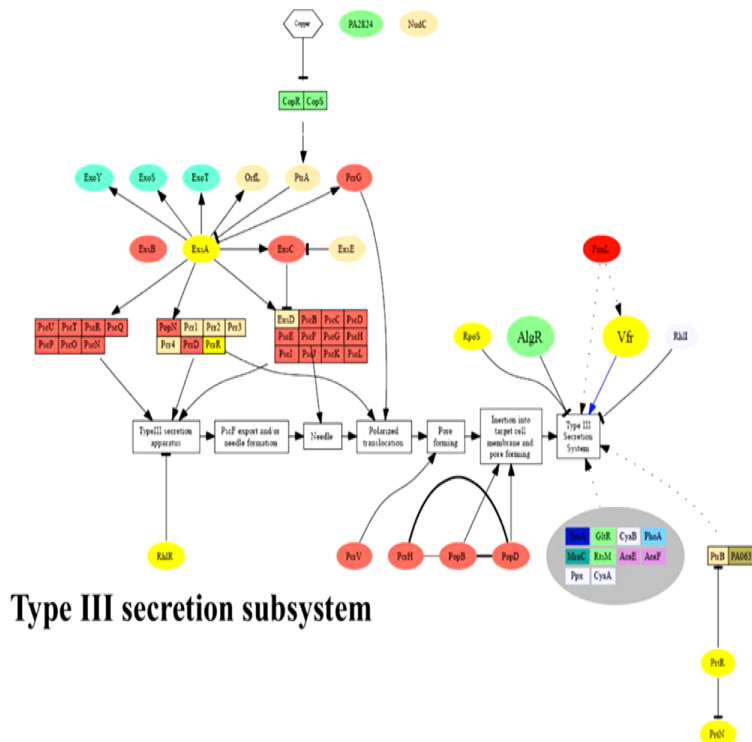
The proposed PHW service will include as content potentially all sorts of traditional EHR (or EMR)⁴ attributes of medical information and lab measurements as standardised in LOINC.⁵ There are currently hundreds of EHR and EMR on the market without any standards. We will favour a template of the OpenEHR in this proposal. In fact, our previous results will help this situation significantly. In Kim et al. (2009), we have developed several indexing and search algorithms for integrating multiple databases semantically. From the database perspective, this task can be formulated concretely and solved. We will use the openEHR as the primary care database format. Certain associated LOINC data will be collected for each patient, while their meanings will be developed as knowledge bases with linkages to various medical ontological frameworks.⁶ A challenge is how to integrate traditional medical information with genomic information?

In Chen (2007a, 2007b), Kim and Chen (2007, 2009), Chen and Kim (2007), Wu et al. (2007), Kim et al. (2009) and Song and Chen (2009), we have text-mined the PUBMED literature into semantic knowledge bases. In the proposed PHW service, PUBMED may be potentially linked, furthermore knowledge may be distilled and gathered to help the users. The indexing of various databases will be managed by the metadata server of Kim et al. (2009), which integrates the semantics of different databases. Although the full problem of semantic integration is very hard, our result in Kim et al. (2009) offers a simpler formulation. Next, we turn to data coming from personalised medicine – genomics, proteomics etc. Our approach is information visualisation in graphical and network models. In Wu et al. (2007), Song and Chen (2009), we have used text mining to extract network models from PUBMED literature, and represent them through information visualisation techniques. In the next section, we will develop automated methods to construct graphical and network models from genomic data sets for personalised medicine. This seems to be a powerful way to serve the users by integrating text mining from literature and computation from data, two different ways of thinking. In this proposal, we will explain only the gene function and their regulation networks, leaving the other proteomics and metabolomics to be studied elsewhere. However, there is still considerable challenge to this approach. From the experimental data, embedded in tens of thousands of published literature, it is difficult for the PHW service to extract a comprehensive view of the gene function and their regulation in the human genome. Furthermore, rapid progress of the research on

different genomic sequences within recent years brings in exponential growth of related literature. We will initiate this research direction with only few diseases and few gene regulatory networks. To help PHW users to keep up with all the new information, a complete system that not only compiles the experimental evidences but also logically integrates the knowledge related to gene function and regulation is much desired in the future.

In recent years, the extraction of knowledge from biological literature has received considerable attention. There are two most used methods to extract biological knowledge: either a statistical method based on co-occurrences of proteins or genes (see next section also), or a rule-based extraction method. Statistical methods are good at locating potential gene–gene or protein–protein interactions. But, they usually cannot provide a clear classification of interaction information. Rule-based relation extraction methods can achieve good precision and recall if the manually developed patterns are good enough. In our paper (Song and Chen, 2009), we provided a framework of a rule-based method with the help of potential informative sentence discovering from literature. Therefore, new templates and rules can be incrementally discovered and supplemented. We have obtained results on a bacterium, but the methods can be used on human genome. Figure 1 is the network for a subsystem of the whole biological system:

Figure 1 Gene regulatory networks, that are essential to certain disease, can be constructed by text mining the relevant literature in the PUBMED digital library of the NLM. Illustrated here is the gene regulatory network of a subsystem of the biological system of *Pseudomonas aeruginosa*, which is a bacterium causing cystic fibrosis (see online version for colours)

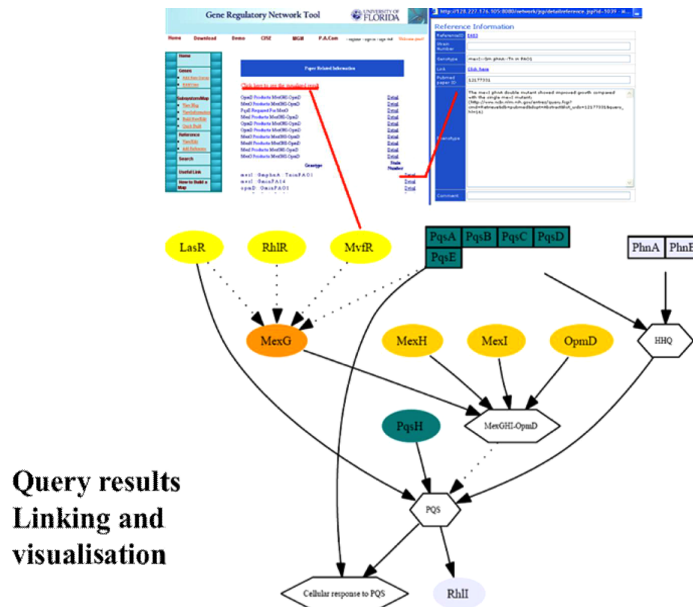


Source: Adopted from Song and Chen (2009)

In Song and Chen (2009), the networks are supplemented by other information visualisation modules: linkages to the PUBMED paper, from which we extracted the network, the genomic database that contains the genes in the network, and other information items. Figure 2 illustrates these and a summary of the knowledge in our information visualisation module. But, this can be improved in several ways. In particular the integration of network models to the traditional medical information in the EHR's and LOINC data will be much desired in clinical and translational sciences. The translational sciences intend to incorporate many new genomic experiments into clinical studies. Our opinion is to start at the basic level of EHR and LOINC information level now. These networks will be indexed with attributes, which can be associated with the database attributes in the EHR and LOINC data sets. We envision some further network models to be constructed to represent these associations and relationships. In a sense, information visualisation plays a vital role in this paper.

As to text mining the genomic literature, fortunately, most biomedical papers provide the gene-strain table in a standard format. We find out that there are some rules to help us to recognise certain tables. Most tables include three columns, the first one for the strain number, the second one for the genotype, and the last one for the reference information. To recognise these columns, we can follow these patterns. The column name for the strain number usually includes strain or plasmid, such as, 'Strain or plasmid', 'Strain, plasmid, or oligonucleotide'. The column name for genotype usually includes genotype or description or characteristics, such as, 'Description or sequence', 'Characteristics' or 'Description'.

Figure 2 The gene regulatory networks (Figure 1) are supplemented by other information visualisation items: linkages to the PUBMED paper, the genomic database that contains the genes in the network, and other information items. These items are followed by a graphical summary of the extracted knowledge in our information visualisation (see online version for colours)



Source: Adopted from Song and Chen (2009)

4 Graphical and network models

The challenge to ‘Clinical and Translational Sciences’ or ‘Personalised Medicine’ rests on the multitudes and volumes of genomic data sets (mostly gene expression profiles). The models are constructed by a statistic data mining methodology, which integrates the probability-theoretic and information-theoretic methods in our recent 2010 BMC Systems Biology paper (Tang et al., 2010). The paper (Tang et al., 2010) has studied few data sets, such as cell cycle and cancer, but we may study many different kinds of clinical cases here. In Tang et al. (2010), we introduced a dimensionless metric based on the above-mentioned concepts in defining pairwise regulatory strengths, and a phase-shift metric in determining pairwise regulatory directions from the digital signal processing perspective. Finally, considering different biological backgrounds and inference purposes, we proposed a multi-objective combinatorial optimisation for the related inference tasks. The framework renders the possibilities of incorporating previously acquired biological knowledge and specific analysis purposes for multiple datasets. Correlation analysis will reveal the strength of a linear relationship between two random variables; statistically correlation or correlation coefficient represents the departure of two random variables from independence. Among several correlation coefficients available, the Pearson product-moment correlation coefficient is applicable to data of diverse characteristics. Normally, the correlation coefficient $\rho_{X,Y}$ is denoted as:

$$\begin{aligned} \rho_{X,Y} &= \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E((X - \mu_X)(Y - \mu_Y))}{\sigma_X \sigma_Y} \\ &= \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}} \end{aligned} \quad (1)$$

where cov indicates covariance, E is for the expected value operator, $\mu_X = E(X)$ and $\sigma_X^2 = E[(X - E(X))^2] = E(X^2) - E^2(X)$. Since the introduction of the correlation coefficient, several scientists also attempted to offer guidelines for the interpretation of the measure. Cohen indicated that the proposed interpretative criteria were in general arbitrary and thus specific treatments should be adopted for specific cases, ranging from physics to social sciences (Tang et al., 2010). Besides the parametric statistic as Pearson’s correlation, these nonparametric correlation metrics as the χ^2 test, Spearman’s ρ , and Kendall’s τ , etc. are much more applicable to statistical problems following diverse non-normal distributions. To quantify the mutual dependence of the two variables, besides the above-mentioned correlation measure, the mutual information is frequently adopted as an alternative in the probability and information theoretic applications. The mutual information of two discrete random variables, X and Y can be defined in a formal term as:

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log \left(\frac{p(x,y)}{p_1(x)p_2(y)} \right) \quad (2)$$

where $p(x,y)$ denotes the joint probability distribution function of X and Y , and $p_1(x)$ and $p_2(y)$ denote the marginal probability distribution functions of X and Y respectively.

The above-mentioned measure normally takes the well-defined form as $I(X, Y, b)$, where b denotes the relevant base. Alternatively, a base of 2 could be specified since the most common unit of measurement of mutual information is the bit. Thus for the analysis

within the following context, we take the base 2 uniformly. The relationship between mutual information and entropies is illustrated as follows:

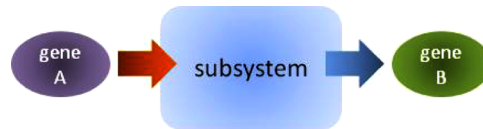
$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) = H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X,Y) \\ &= H(X,Y) - H(X|Y) - H(Y|X) \end{aligned} \quad (3)$$

where $H(X)$ and $H(Y)$ denote marginal entropies, $H(X|Y)$ and $H(Y|X)$ the conditional entropies, and $H(X,Y)$ the joint entropy between the two variables, X and Y . The non-negativity and symmetry also characterise the mutual information measure, i.e., $I(X;Y) \geq 0$ and $I(X;Y) = I(Y;X)$. The above-mentioned measures illustrated the correlation and dependence relationships between two random variables within a system. Normally, these random variables represent nodes within a graphical structure. And, the internuncial weights may be denoted by the probability of association between these pairs under investigation. Since both mutual information and correlation coefficient are dimensionless metrics for vector quantities, here we introduce a uniform associativity measure for illuminating associativity between candidate gene pairs. Within this uniform associativity measure, the quantities of mutual information and correlation metrics can be projected onto the real and imaginary axes of a two-dimensional geometric plane respectively. The metric is given as,

$$\begin{aligned} AM_i &= w_{i1} \overline{MI_i} + w_{i2} \overline{Cor_i} = [w_{i1} MI_i] + j[w_{i2} Cor_i] = |AM_i| \angle \alpha_i \\ &= \sqrt{[w_{i1} MI_i]^2 + [w_{i2} Cor_i]^2} \angle \tan^{-1} \left(\frac{w_{i2} Cor_i}{w_{i1} MI_i} \right), \quad i \in N \end{aligned} \quad (4)$$

where MI_i and Cor_i denote the mutual information and correlation quantities, respectively w_{i1} and w_{i2} the weights for both quantities, and α_i is the phase difference for the i th pair candidate. Note that the weights here aim to leverage the possible asymmetric distribution within the datasets of the above-mentioned subterms, MI_i and Cor_i . These weights can be derived according to some previously acquired knowledge or specific theoretical hypothesis, e.g., the respective centroids of the subterms' datasets. For example, the genomic data can be gene expression profiles. They are mostly discrete time-series data sets. The data samples are diverse expression densities measured at multiple time points, and the data intervals represent sampling periods. Then measurements of N gene samples are compared against each other, resulting in totally $N(N-1)$ pairwise comparisons. One can utilise signal processing to clustering and comparing the similarity of time-series gene expression. For each potential regulatory role between two genes from a large regulatory network, the investigated genes and their activities can be modularised as a subsystem. The expression patterns of the two genes might be viewed as input and output signals respectively, illustrated in Figure 3.

Figure 3 Each pairwise gene associations can be viewed as a subsystem with gene expression patterns as input and output signals respectively (see online version for colours)



For each gene pairs, we calculate the coherence, gain and phase shift from the associated discrete Fourier transform (DFT) on the input and output signals. Coherence of signals a and b is a function of the power spectral density (PSD) and the cross power spectral density (CPSD), which is defined as,

$$\text{Coh}_{ab}(f) = \frac{|\text{CPSD}_{ab}(f)|^2}{\text{PSD}_{aa}(f) \cdot \text{PSD}_{bb}(f)} \quad (5)$$

where $\text{PSD}_{aa}(f)$, $\text{PSD}_{bb}(f)$ and $\text{CPSD}_{ab}(f)$ measure the power spectral density and cross power spectral density of the associated pairwise signals, respectively. The symbol f represents a frequency-domain metric. Normally, signals a and b should be of the same length. A coherence of 1 represents a scalar-multiples relationship between two investigated signals, while 0 defines such a relationship of being not linearly related. The Transfer Function (TF) between two associated input-output signals is to measure the signal amplification and related time lag or latency properties, which is defined as follows,

$$\text{TF}_{ab}(f) = \frac{\text{PSD}_{ab}(f)}{\text{PSD}_{aa}(f)} \quad (6)$$

where the $\text{PSD}_{ab}(f)$ and $\text{PSD}_{aa}(f)$ remain the same defined as earlier. The regular TFs will be of the complex-valued form, the arctangent of which are the corresponding Transfer Phases (TPs), and the absolute values derive the related Transfer Gains (TGs), represented respectively as follows,

$$\text{TP}_{ab}(f) = \arctan \left[\frac{\text{PSD}_{ab}(f)}{\text{PSD}_{aa}(f)} \right] \quad (7)$$

$$\text{TG}_{ab}(f) = \text{abs} \left[\frac{\text{PSD}_{ab}(f)}{\text{PSD}_{aa}(f)} \right]. \quad (8)$$

Theoretically, the TP illustrates the phase shift between the investigated pairwise signals, i.e., the input and output. The bigger the ratio is, the less energies the output will lose. An effective evaluation criterion for those metric functions is the related coherence value, i.e., at such frequencies where coherence values are higher, the corresponding TPs and TGs are much more trustable than those at other frequencies. Through the above-mentioned methods, regulatory directions may be readily derived from gene pairs under investigation. The advantages of such a signal processing metric cover the following aspects:

- Due to diverse regulatory durations, intensities and even initial conditions and environments, the time-variant properties are the major characteristics within most regulatory processes, thus time-series expression profiles might measure multiple genetic regulatory processes and capture inherent major characteristics simultaneously. The above-mentioned multiple frequency metrics try deciphering those versatile regulatory properties in a quantitative way.
- TP shifts may reveal underlying changes of regulatory direction and strength at different frequencies.

Such specificities as time-varying regulatory direction and strength are relatively common among the genetic, metabolic and protein–protein interaction networks during developmental and evolutionary processes. Furthermore, the TG at multiple frequencies acts as a concrete indicator measuring different possibilities of network connectivities and variations of regulatory information. Such kinds of flexible network connectivities and regulatory information characterise major genetic regulatory processes from network and information-theoretic perspectives. Avoiding a lot of technical details, we illustrate the results of Tang et al. (2010) (see details there) in Figures 4 and 5.

Figure 4 The global phase-shift statistics distribution for the cell cycle regulatory network (totally 83 pairwise candidates). The phase-shift statistics vary as functions of the gain thresholds. The blue bold curve represents the integral tendency of gene pairs with leading phase shifts (positive), the red for the pairs with lagging phase shifts (negative), and the green for those without detected phase shift (undirected), i.e., there might be no regulatory activities between corresponding gene pairs (see online version for colours)

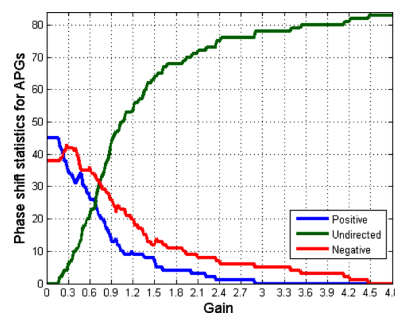
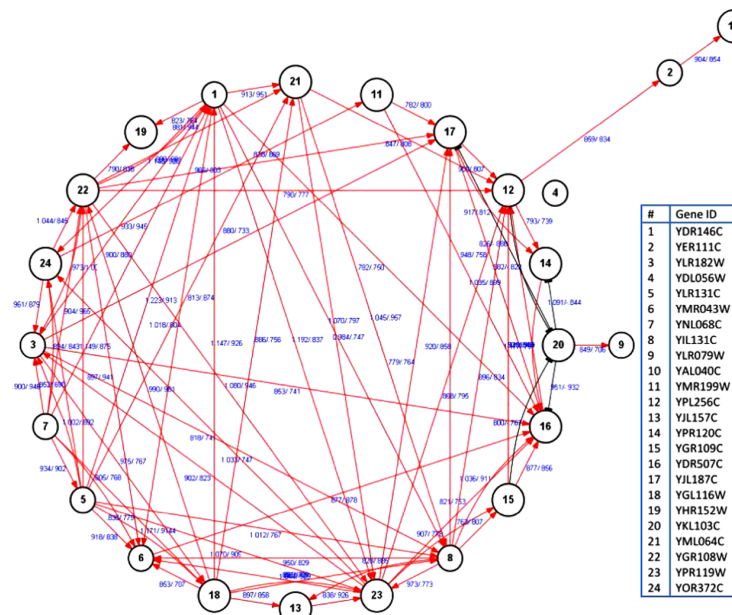


Figure 5 The hypothetic interweaved cell cycle regulatory network constructed based on the MIONA framework. Each gene/protein is denoted as a black-edged circle. The calculated associativity metric and phase-shift information between gene pairs are marked as blue along each bilateral link (see online version for colours)



The above-mentioned methodology provides a general approach to extract gene regulatory networks and medical pathways for constructing personalised medicine profiles to be stored in the SHW data centre services. The next section describes how to access and utilise such information by patients and healthcare workers using wireless communication devices.

5 Conclusion

This paper provides a practical and feasible approach to personalised medicine in the framework of clinical and translational sciences. Currently, we are seeking collaborations for experimenting this system design and methodology.

References

- Chen, S. (2007a) 'Fusion of multimedia information in biomedicine, advances in multimedia information processing – PCM 2007', *8th Pacific Rim Conference on Multimedia*, Hong Kong, China, 11–14 December, Proceedings, Lecture Notes in Computer Science, 4810 Springer, pp.494–450.
- Chen, S. (2007b) 'Digital preservation: organizational commitment, archival stability, and technology continuity', *Journal of Organizational Computing and Electronic Commerce*, Vol. 17, No. 3, August 24, Taylor & Francis, pp.205–215.
- Chen, S. and Kim, H. (2007) 'Automated-linking PUBMED documents with GO terms using SVM', *Journal of Data Science*, Vol. 5, pp.1–9.
- Chen, S., Choo, C. and Chow, R. (2006) 'Internet security: a novel role/object-based access control for digital libraries', *Journal of Organizational Computing and Electronic Commerce*, Vol. 16, No. 2, pp.87–103.
- Kim, H. and Chen, S. (2007) 'Ontology search and text mining of MEDLINE database', in Pardalos, P.M., Boginski, V.L. and Vazacopoulos, A. (Eds.): *Data Mining in Biomedicine*, Springer, pp.177–193.
- Kim, H. and Chen, S. (2009) 'Automated ontology linking by associative Naive Bayesian classifier', *Pattern Recognition*, Vol. 42, No. 9, pp.1777–1785.
- Kim, H., Choo, C. and Chen, S. (2009) 'Mediation of federated intelligent information systems by a meta digital library server', *Journal of Intelligent Information Systems – The Journal of Intelligent*, Springer-Verlag, ISSN: 0925-9902 (print version), ISSN: 1573-7675 (electronic version), www.springer.com/computer/security+and.../10844.
- Song, Y. and Chen, S. (2009) 'Text mining biomedical literature for constructing gene regulatory network', *Interdisciplinary Sciences – Computational Life Sciences*, Vol. 1, No. 3, pp.179–186.
- Tang, B., Wu, X., Tan, G., Chen, S., Jing, Q. and Shen, B. (2010) 'Computational inference and analysis of genetic regulatory networks via a supervised combinatorial-optimization patterns', *BMC Systems Biology*, Vol. 4, Suppl. 2, p.S3, <http://www.biomedcentral.com/1752-0509/4?issue=S2/S3>
- Wu, W., Song, Y., Jin, S. and Chen, S. (2007) 'An interactive map of regulatory networks of *Pseudomonas aeruginosa* genome', in Eskin, E. *et al.* (Eds.): *RECOMB 2005 Workshop on Regulatory Genomics*, LNBI 4023, Springer-Verlag, Berlin Heidelberg, pp.1–10.

Notes

¹<https://www.ctsi.ufl.edu/>

²<http://www.hhs.gov/ocr/privacy/>

³J. Barkley, T. Cincotta, S. Gavrilu, RBAC/Web Release 1.1, May 1998, <http://hissa.ncsl.nist.gov/rbac/>

⁴<http://www.openehr.org/home.html>

⁵<http://loinc.org/>

⁶For examples, SNOMED CT, UMLS and others.