Big-data-driven Performance Analysis, Prediction and Control of Smart Factory

Presenters: Hong Yan (EE)
Min Xie (SEEM, SDSC)
Ding Zhang (SEEM)

Outline

Introduction to smart factories

Abnormity diagnosis

Performance prediction

Data-driven control

Other projects on AI and data analysis

Project Team

CityU
- Hong Yan (EE, Project Leader)
- Ray Cheung (EE)
- Sam Kwong (CS)
- Min Xie (SEEM, SDSC)
- Moshe Zukerman (EE)

GDUT Collaborators
- Qing Liu (Project Leader)
- Jiewu Leung
- Ding Zhang

PhD Students
- Tahir Mahmood
- Feng Zhu

Smart Factory

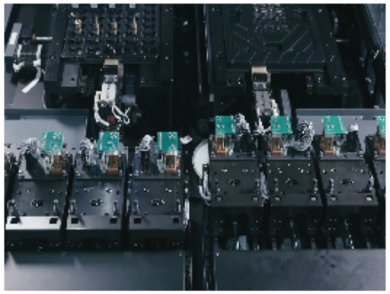
The explosion of manufacturing data

Automation → Information → Intelligence

Categories: Product-Process-Equipment-Management

> The trend of using efficient machines instead of labors

> The requirement of high-speed, high precision and high-reliability



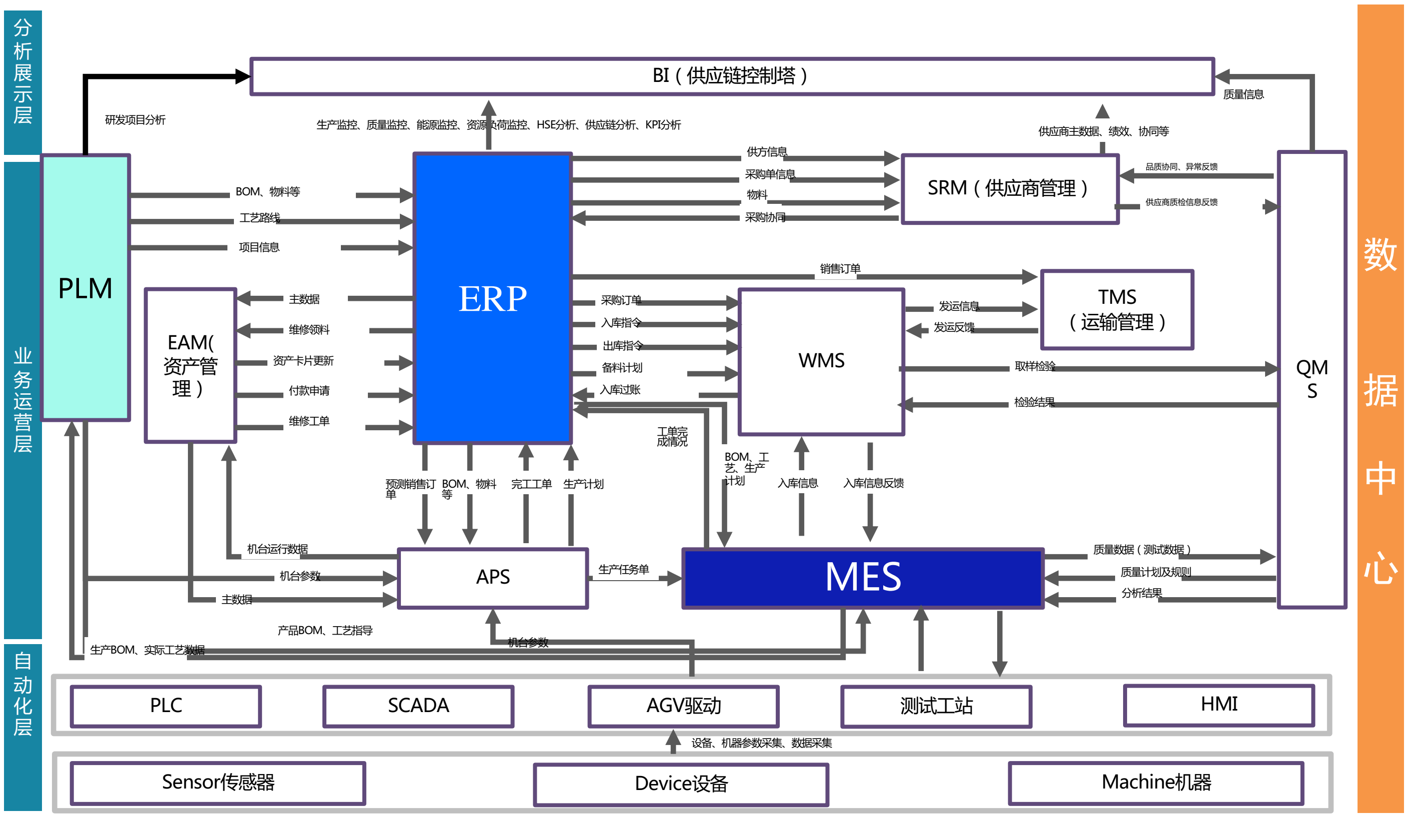
References: https://www.ioranges.cn/
http://www.moqie.com/ArticleView_13765.aspx

香港城市大學
City University of Hong Kong

The complexity of data driven accurate decisions

Multiple sources-multiple dimension-structure/unstructured data

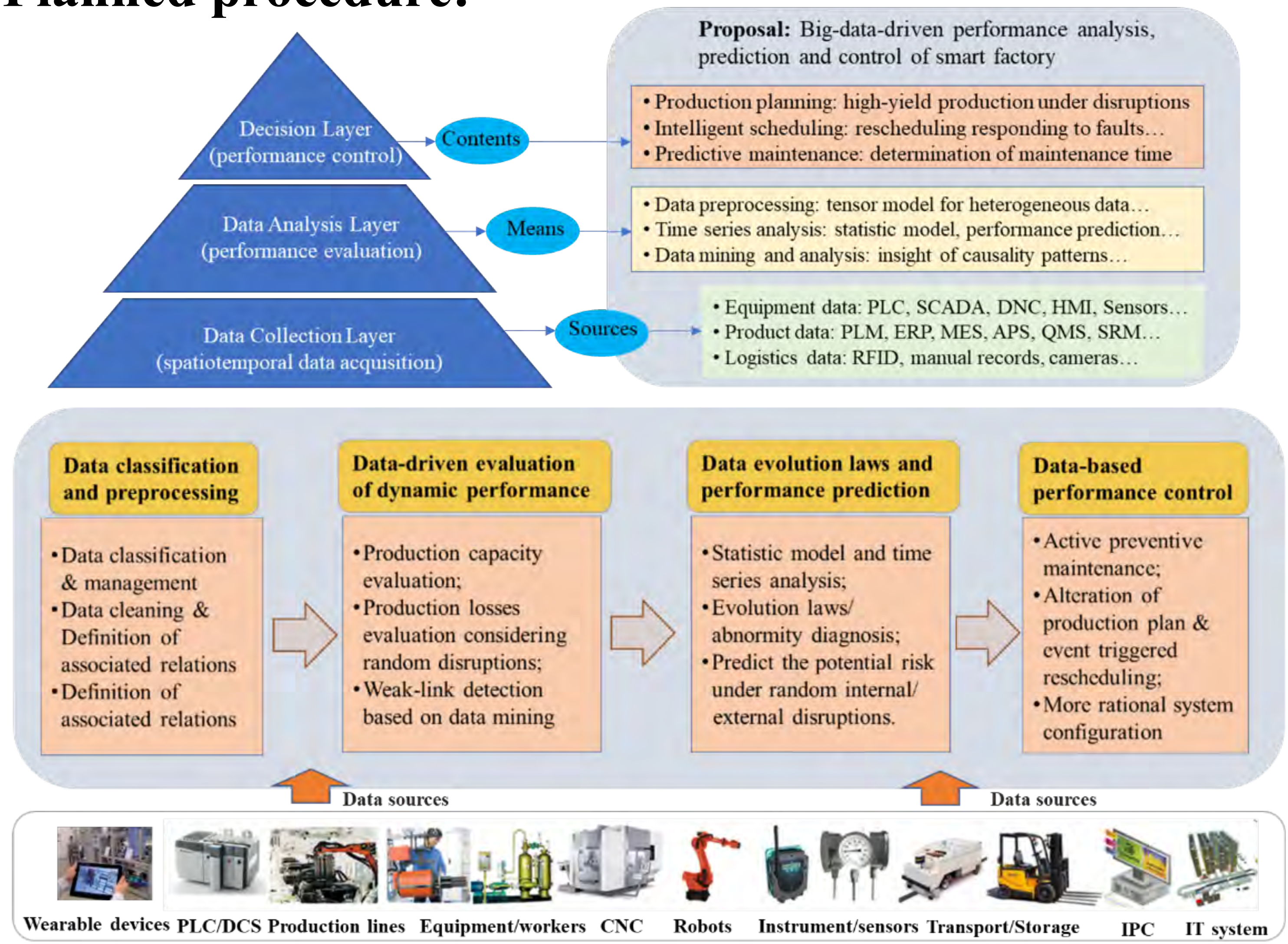
Information-based systems: PLM/ERP/MES/APS/QMS/SRM…



Product Lifecycle Management (PLM)
Enterprise Resource Planning (ERP)
Manufacturing Execution System (MES)
Customer Relationship Management (CRM)
Advanced Planning and Scheduling (APS)

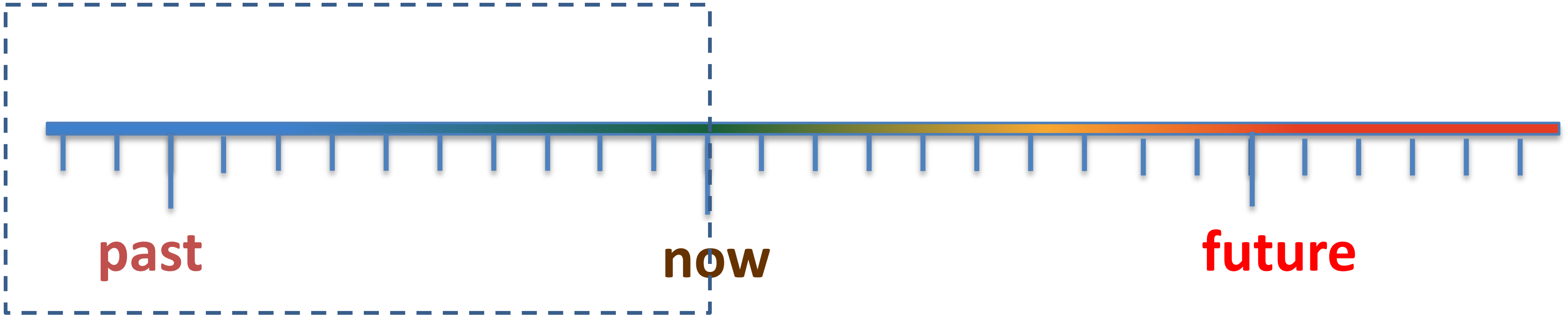
Business Intelligence (BI)
Programmable Logic Controller (PLC)
Supervisory Control and Data Acquisition (SCADA)
Quality Management System (QMS)
Warehouse Management System (WMS)

Planned procedure:



State Awareness & Abnormity Diagnosis

(Performance Evaluation)

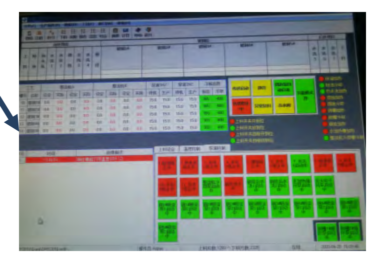


To achieve the state awareness / synchronization,

- Build a **Digital Twin factory** with state replication from the physical system (engineering problems)
- Assistant abnormity diagnosis and alarm (academic problems)

Visualization patrol of PCB factory

Data management



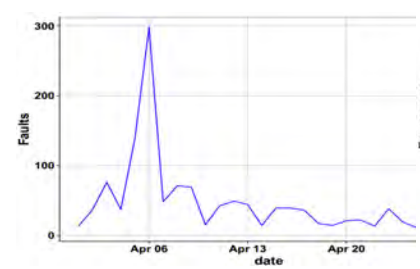
The efficient interconnection of data is the basis for further performance control

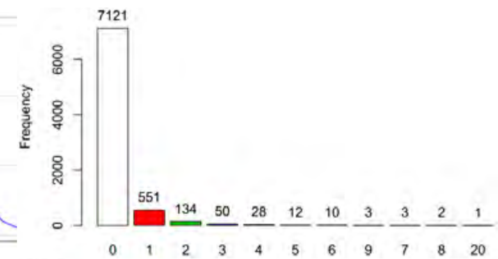
Reference: http://www.korbe.com.cn/productshow_51.html

Typical application: electronic plating line.

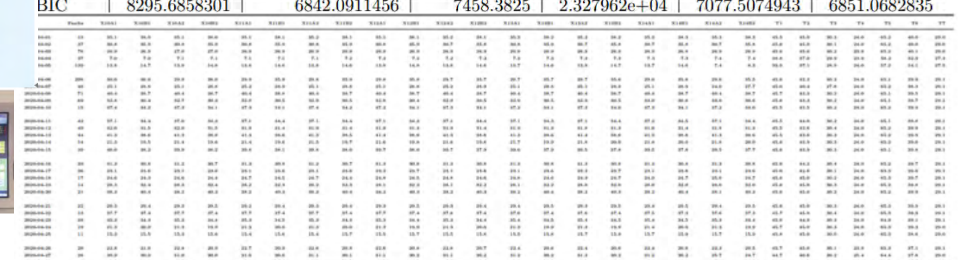
- Data acquisition: multiple sensors including current data of main processes, the temperature data of plating tanks, and other production order data.
- Data preprocessing: improve the data quality.
- Statistic analysis: distribution fitting, statistic process control …

Product and production line



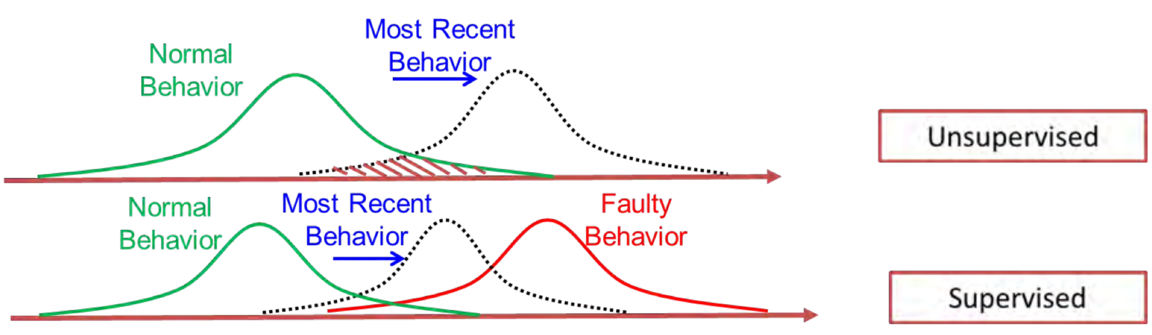


	Poisson	Negative Binomial	COM-Poisson	Bell	ZIP	ZINB
par1	0.1623500	0.1623735	0.1397	1.409993e-01	0.0532747	-1.8179183
SE.par1	0.0045288	0.0068002	0.0036	3.682300e-03	0.0422319	0.0421333
par2	NA	0.1294874	0.0000	NA	NA	-2.0440469
SE.par2	NA	0.0085820	0.0001	NA	NA	0.0665011
par3	NA	NA	NA	NA	1.7041257	-9.4046001
SE.par3	NA	NA	NA	NA	0.0487290	56.1143120
Loklik	-4143.3546576	-3412.0690578	-3720.2147	-1.163084e+04	-3529.7772322	-3412.0693693
AIC	8288.7093152	6828.1381156	7444.4294	2.326367e+04	7063.5544643	6830.1387386
BIC	8295.6858301	6842.0911456	7458.3825	2.327962e+04	7077.5074943	6851.0682835

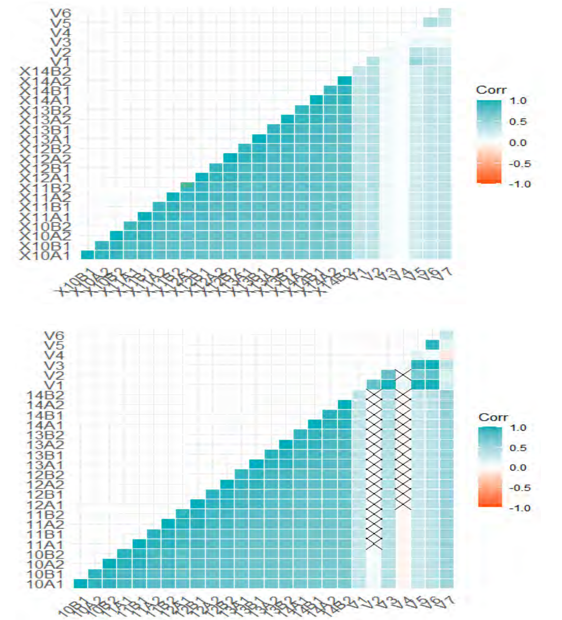


Typical application: electronic plating line.

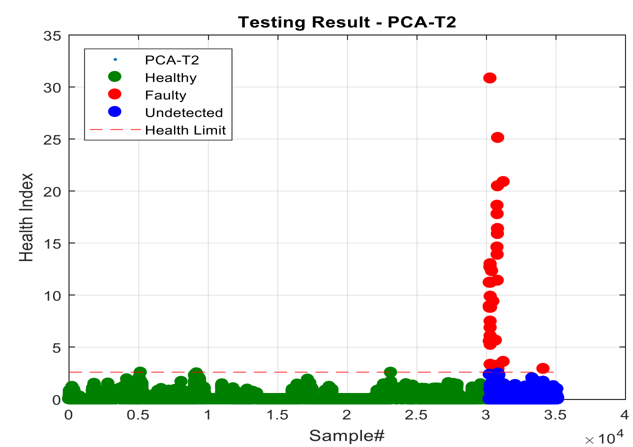
- Correlation Analysis: define the relationship between each sensor variables and select the important sensor variables.
- Fault Detection: based on the data-driven model and machine learning algorithms, build a baseline models for fault detection, which can be referred as Health Assessment.
- Fault Diagnosis: after fault detection, identify sensor variables that are relevant to the occurrence of faults, in order to track down the root cause of the fault.



Data-driven health assessment

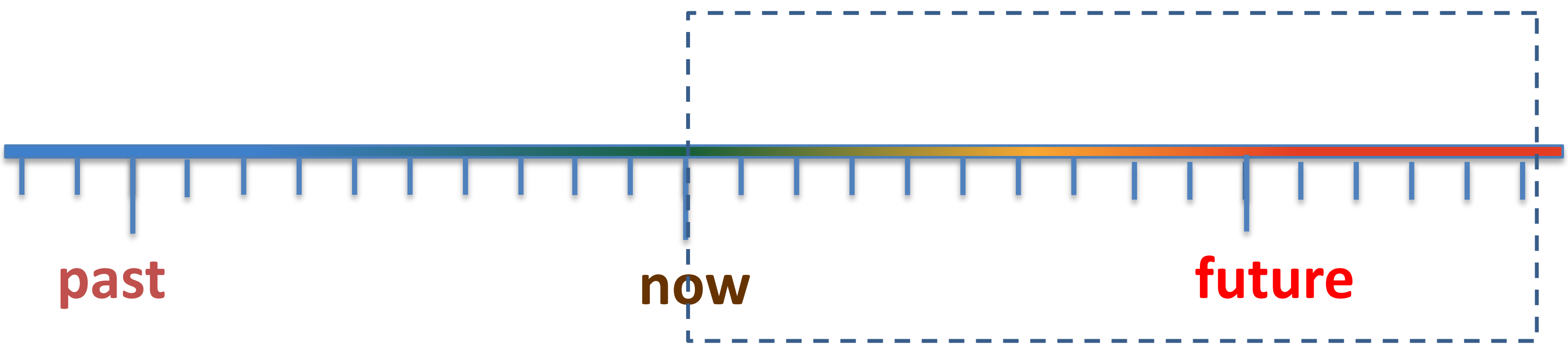


The correlation coefficient between the sensor variables



Fault detection results based on baseline model

Performance Prediction & Data Evolution Laws

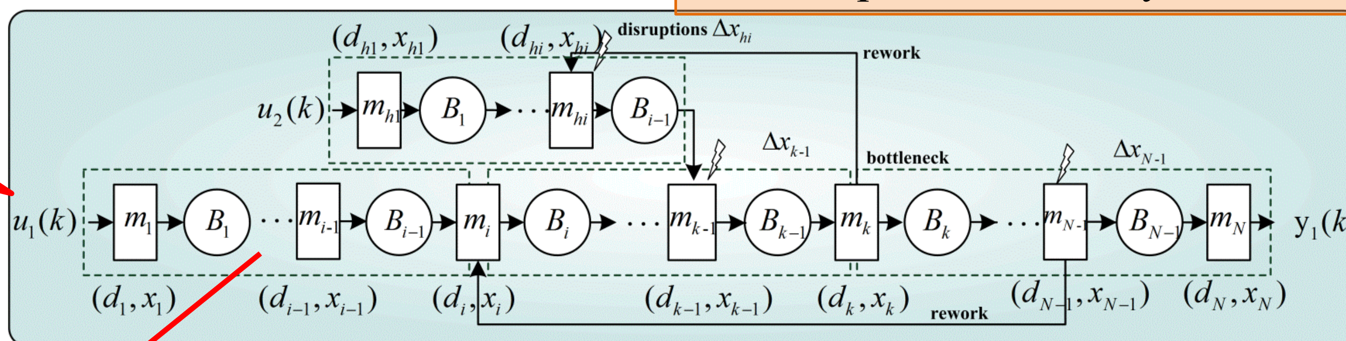


Application 1: performance prediction under disruptions (model driven approach)

Establish the basic production model based on the state awareness (abnormities, disruptive events such as faults, quality flaws, shortage of materials, process postpone……)

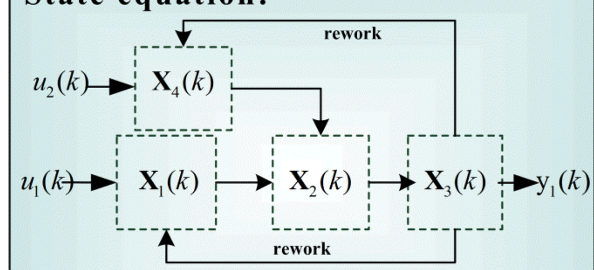
A smart phone assembly case

Monitoring disruption data (or predictive disruptions)



Main manufacturing process

State equation:



$$\mathbf{X}(k+1) = \mathbf{A}_k \mathbf{X}(k) \oplus \mathbf{B}_k \mathbf{U}(k+1)$$

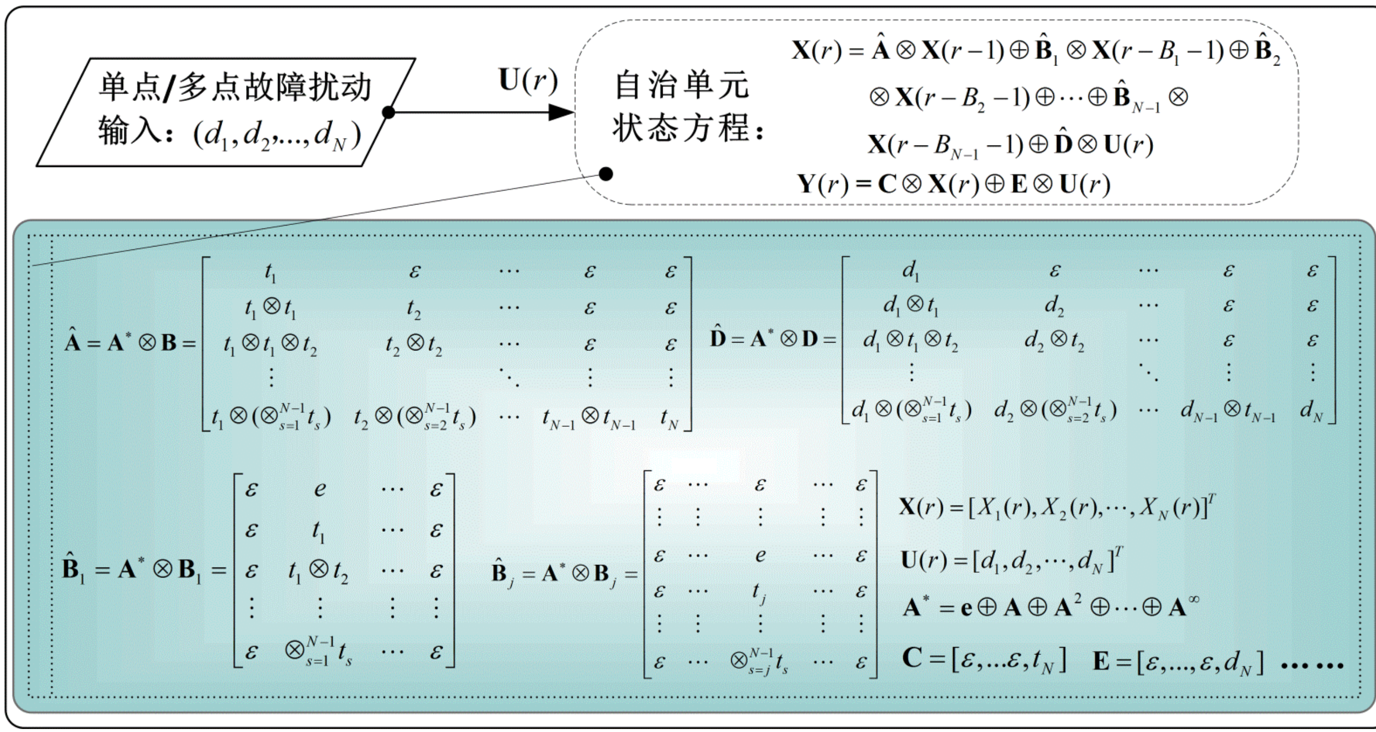
$$\mathbf{Y}(k) = \mathbf{C}_k \mathbf{X}(k)$$

What are the system responses under disruptions?

Application 1: performance prediction under disruptions (model driven approach)

Predict the performance evolution under disruption inputs
(discrete manufacturing dynamic system, delivery delay, customer satisfaction degree……)

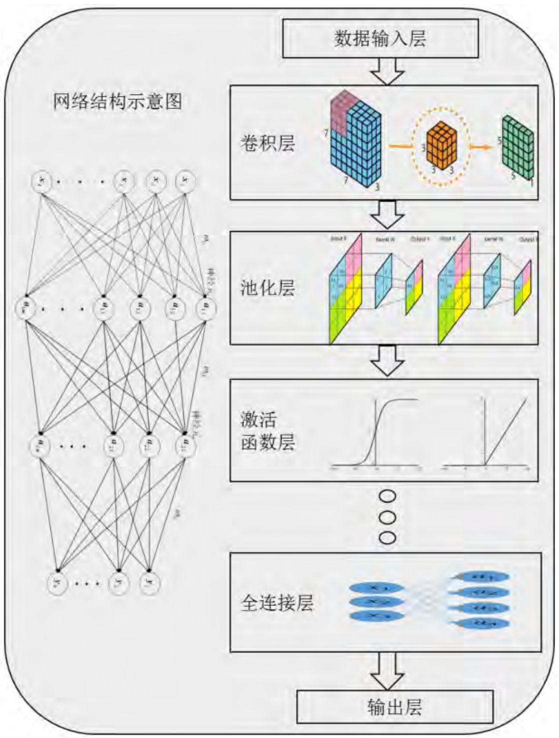
Check whether the disruption attack reach the safety margin.



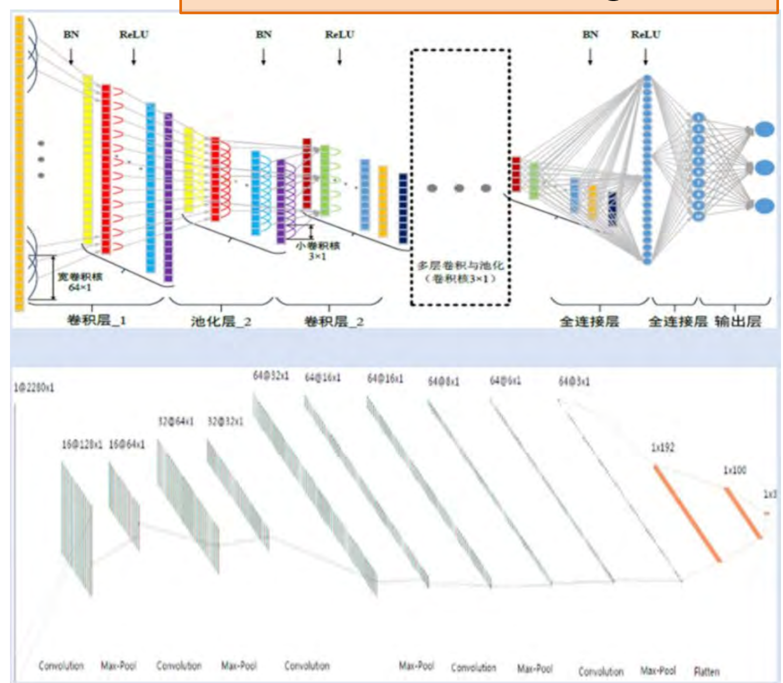
Reference: D Zhang, M Xie, Q Liu, J Leng, and H Yan, paper submitted.

Application 2: performance prediction under history data (data driven approach)

Predict the due-time reliability under history order data, with the aim to decide whether to receive the orders or not according to the current production capacity (manufacturing cost, delivery time, ……)

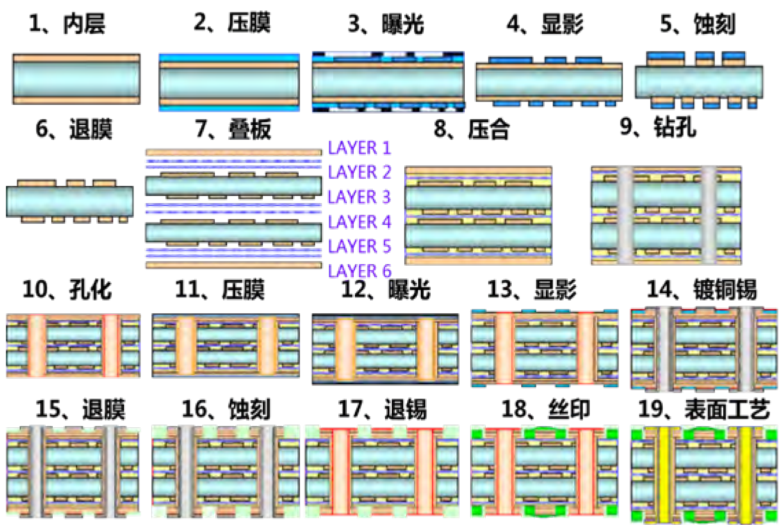


A PCB manufacturing case

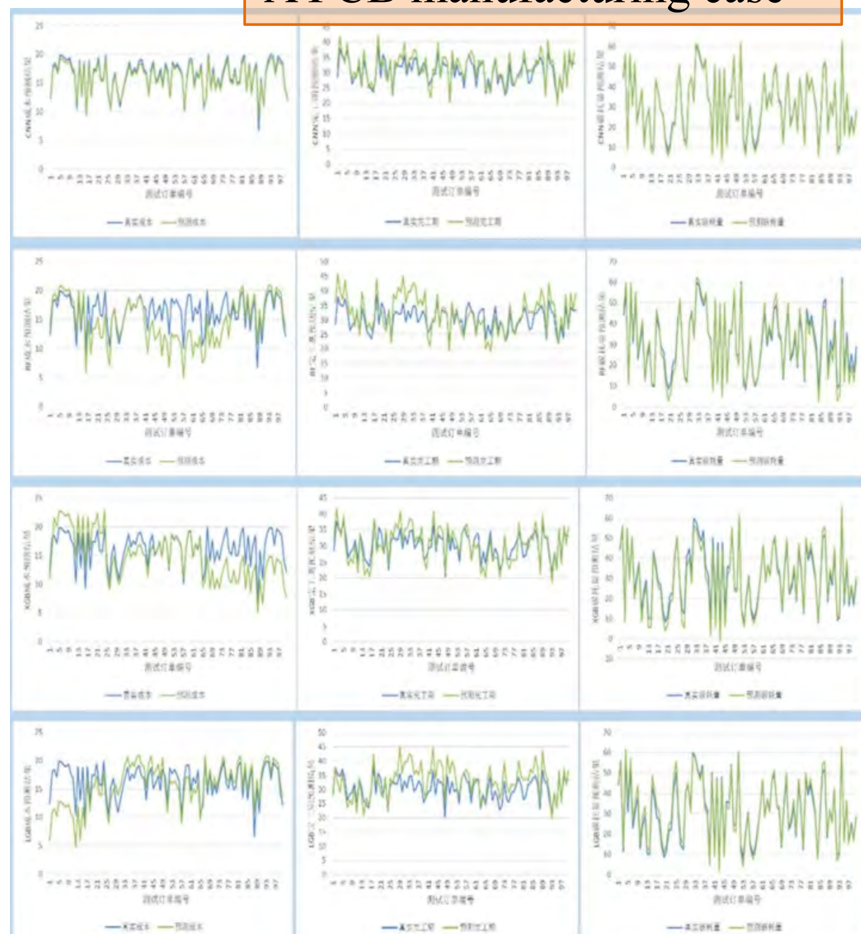


CityU
香港城市大學
City University of Hong Kong

Application 2: performance prediction under history data (data driven approach)



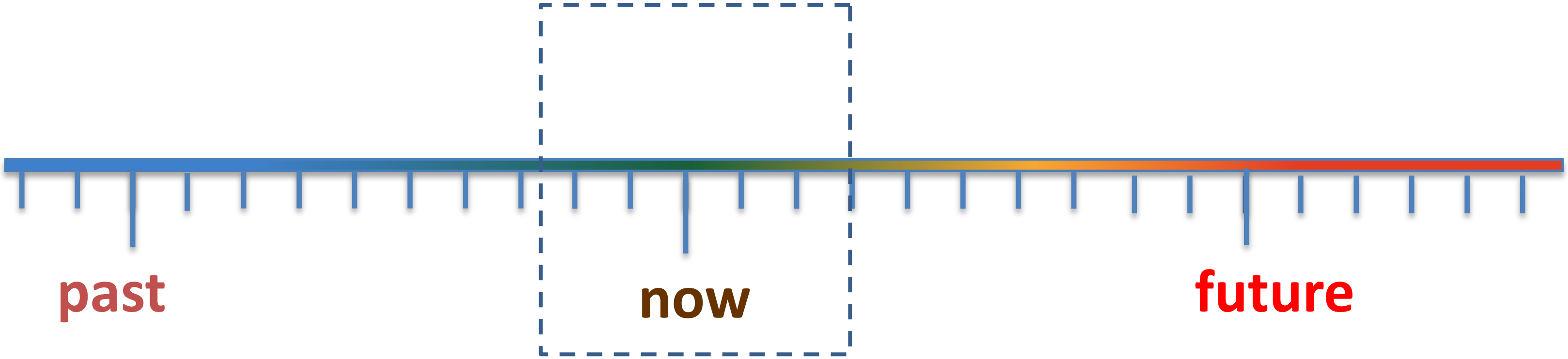
A PCB manufacturing case



Check whether the factory can achieve the due-time delivery within permitted cost.

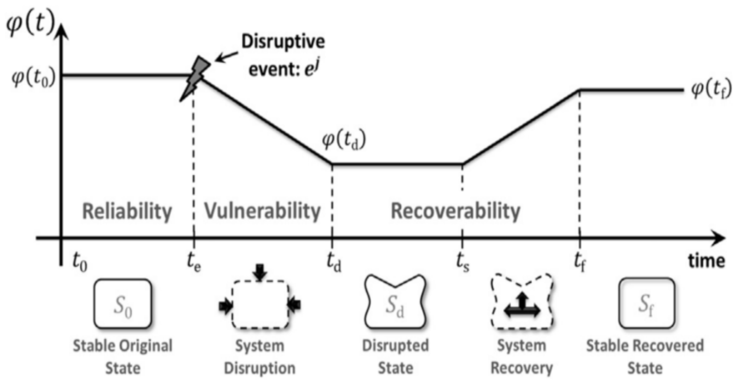
香港城市大學
City University of Hong Kong

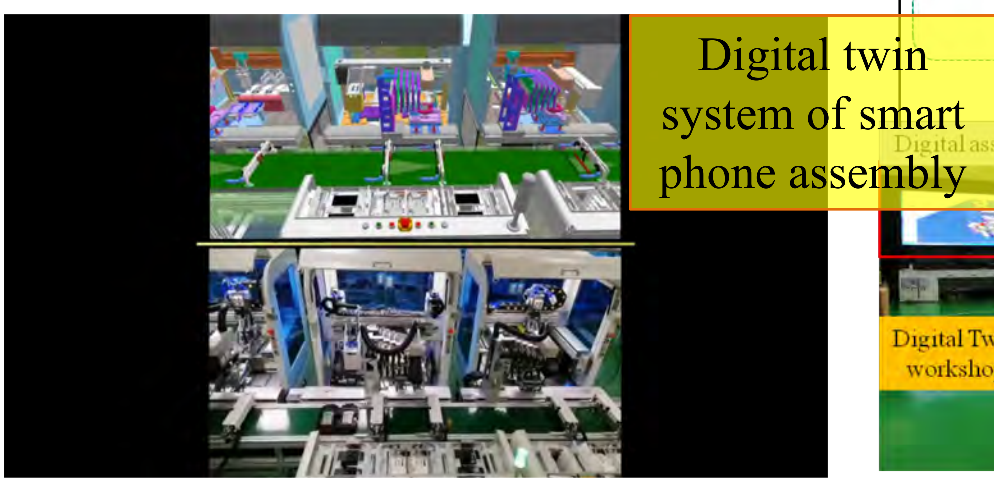
Data-driven Performance Control

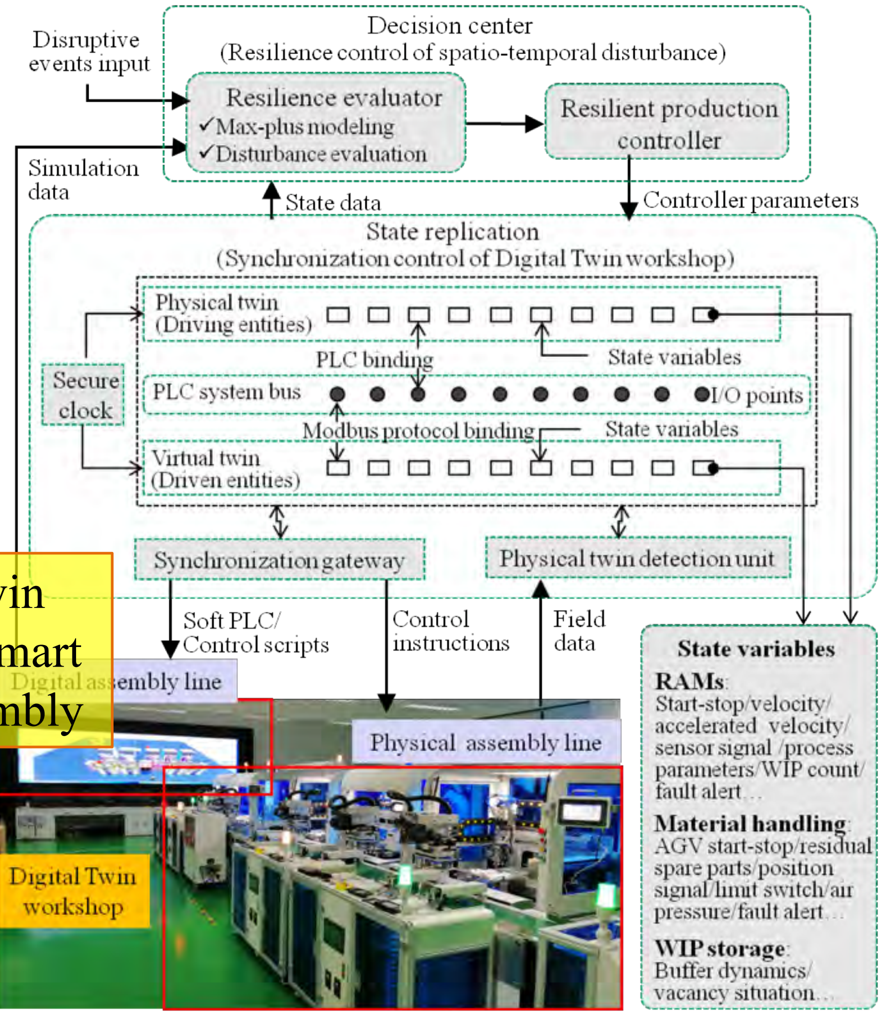


Application: resilience control under disruptions in a digital twin framework

Develop the digital twin platform:

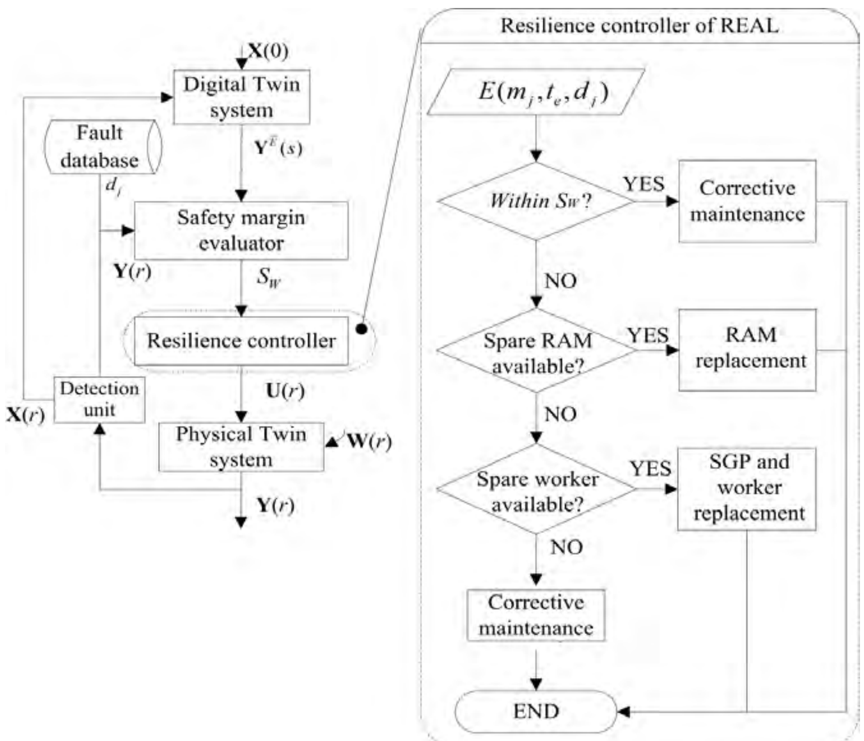


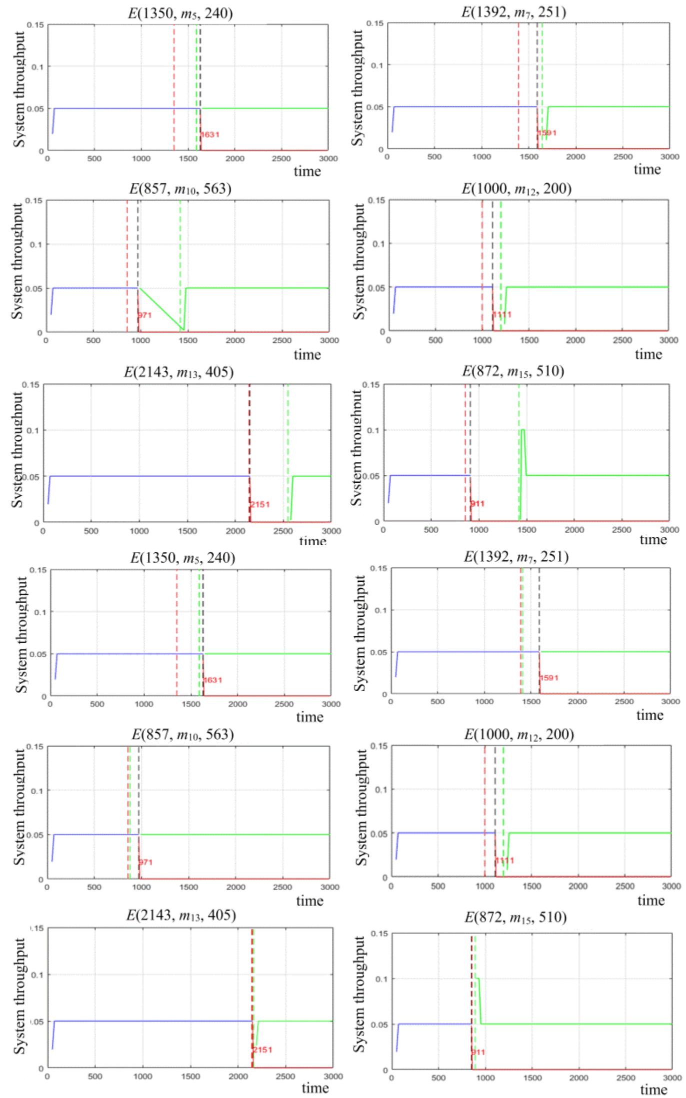




Application: resilience control under disruptions in a digital twin framework

In order to achieve the profit maximization, adaptively control and adjust the production scheme according to real-time data, typically, perform a rescheduling or reconfiguration plan when a fault occurs.





Future Work

- Data collection and pre-processing
- Data analysis → abnormity diagnose + disruption alarm
- Dynamics modeling and decision making

Other Projects on AI and Data Analysis

Theories:
Tensors, Hypergraphs

Algorithms:
Clustering, Co-clustering, Machine learning, Pattern matching

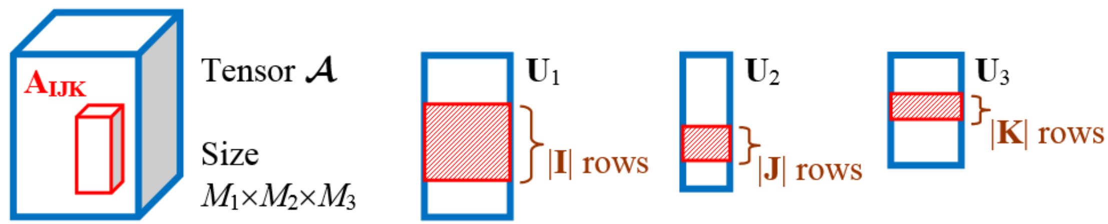
Hardware:
GPU, FPGA, ASIC, Portable electronic devices

Applications:
Human face and facial expression recognition,
Human face animation, Digital entertainment,
Image matching and retrieval, 3D cell imaging,
Object detection/tracking/classification in videos,
Genomic and biomolecular data analysis

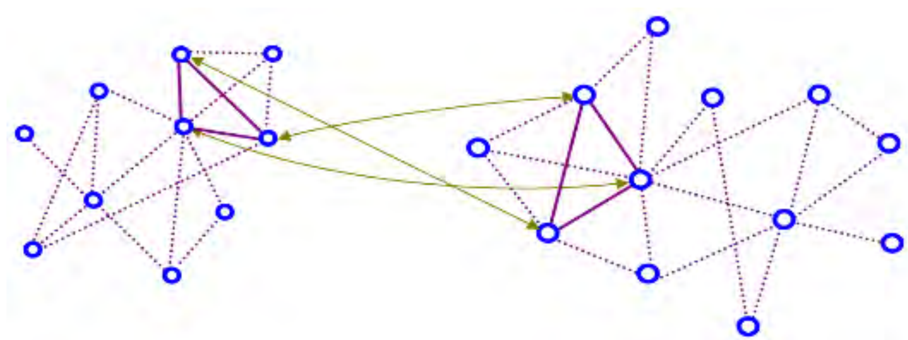
香港城市大學
City University of Hong Kong

Tensor Computing

Co-clustering

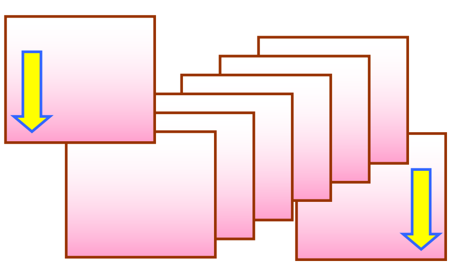


Hypergraph matching



Collaborations:
- CityU
- CUHK
- HKBU
- HKU
- PolyU

Incremental SVD



References:

H Yan, *IEEE System, Man and Cybernetics Magazine*, 23-30, April, 2017.

S Khan, G Xu, R Chan, H Yan, *Expert Systems With Applications*, 90:427-438, 2017.

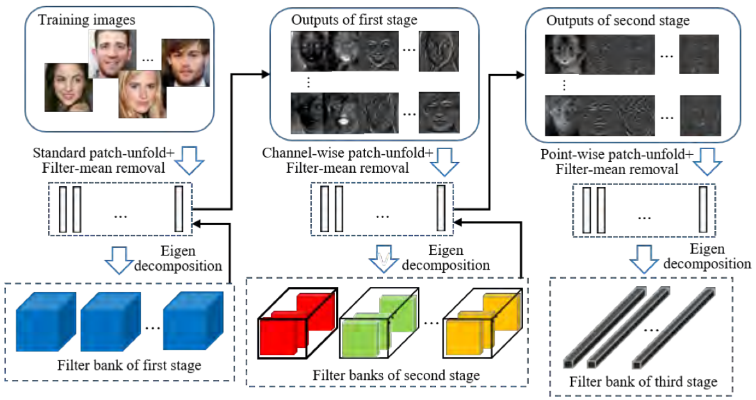
C Cui, Q Li, L Qi, H Yan, *Journal of Global Optimization*, 70(1):237-259, 2018.

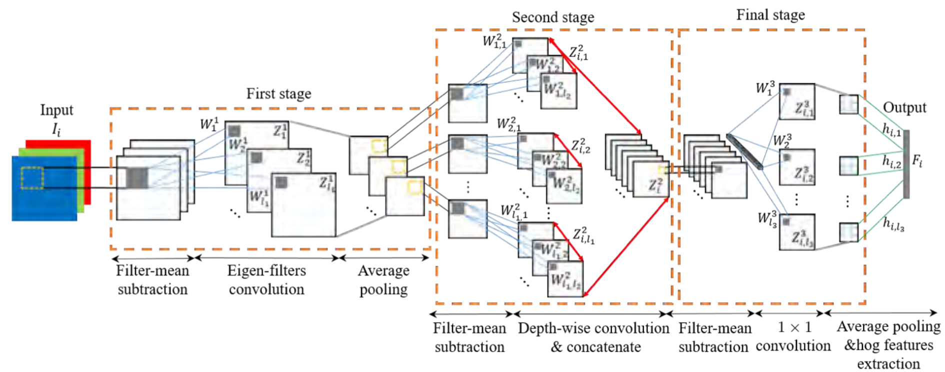
S Khan, M Nawaz, G Xu, and H Yan, *IEEE TCSVT*, in press.

H Zhu, C Cui, L Deng, R Cheung, and H Yan, *IEEE T Cybernetics*, in press.

香港城市大學
City University of Hong Kong

Deep Learning





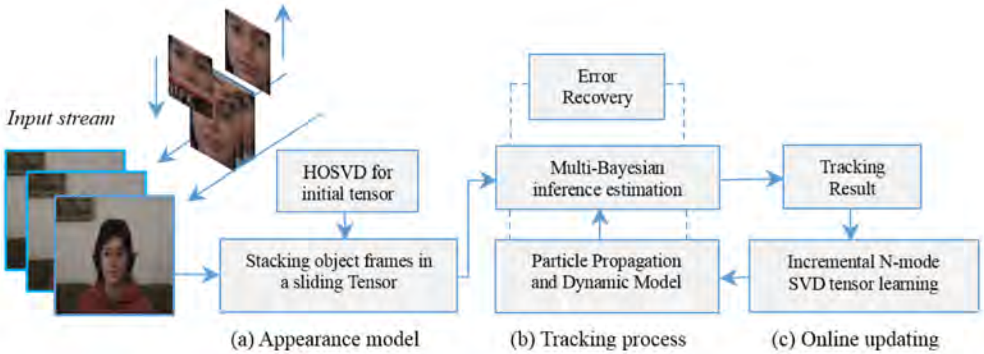
References: X Zhe, S Chen, and H Yan, *Pattern Recognition*, 93:113-123, 2019.

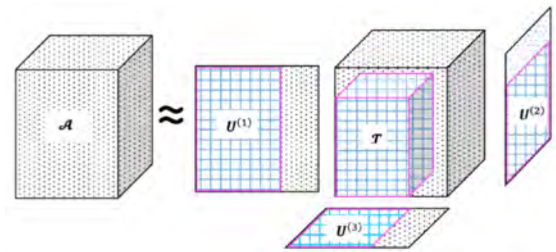
M Zhang, S Khan, and H Yan, *Patter Recognition*, 100:107176, 2020 .

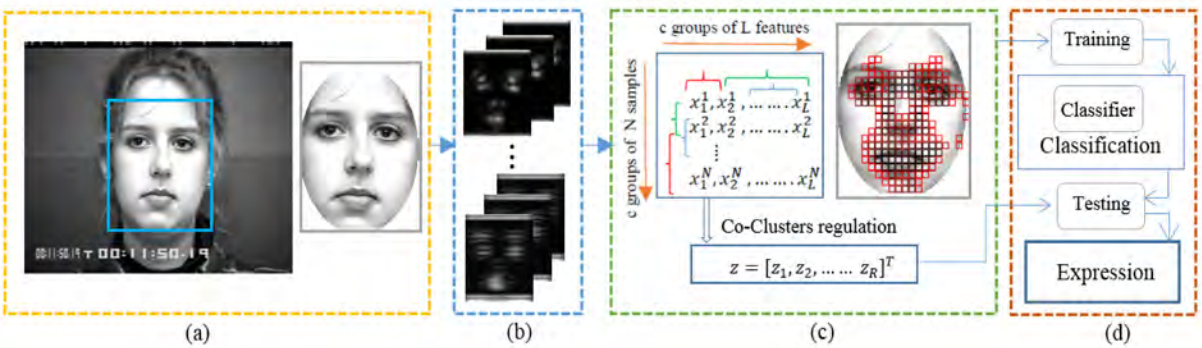
X Zhe, S Chen, and H Yan, IEEE Trans. *Neural Networks and Learning Systems*, 31(5): 1681-1695, 2020.

X Fan, M Jiang, and H Yan. arXiv:2005.03950, 2020.

Facial Expression Recognition







Face tracking **Wavelet features** **Co-clustering** **Recognition**

$$G_{\vec{k}}(\vec{r}\,) = G_{\vec{k},+}(\vec{r}\,) + iG_{\vec{k},-}(\vec{r}\,)$$

$$G_{\vec{k},+}(\vec{r}\,) = \frac{k^2}{\delta^2}\exp\left(\frac{k^2\|r - r_o\|^{\;2}}{-2\delta^2}\right)\cos[\vec{k}(\vec{r} - \vec{r_o}\,)]$$

$$G_{\vec{k},-}(\vec{r}\,) = \frac{k^2}{\delta^2}\exp\left(\frac{k^2\|r - r_o\|^{\;2}}{-2\delta^2}\right)\sin[\vec{k}(\vec{r} - \vec{r_o}\,)]$$

Action Units

Reference: A Amin, H Yan, *Int'l J. Pattern Recognition and Artificial Intelligence*, 23(3): 401-431, 2009.

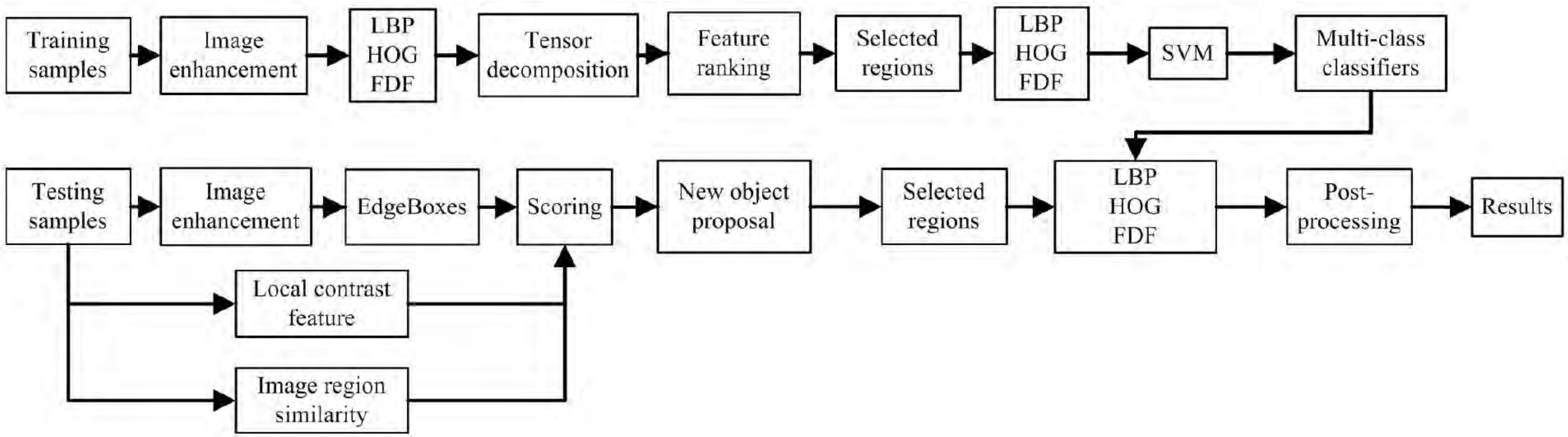
S Khan, G Xu, R Chan, H Yan, *Expert Systems With Applications*, 90:427-438, 2017.

S Khan, L Chen, X Zhe, H Yan, *IEEE Trans. Affective Computing*, in press.

Our contributions:

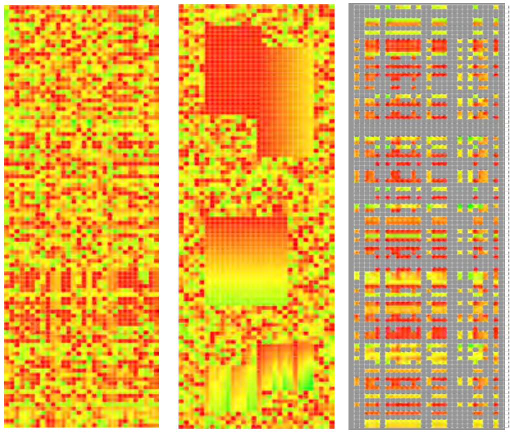
* Online tensor learning
* Low rank representation
* Real-time face tracking
* Feature selection based on co-clustering
* Expression recognition

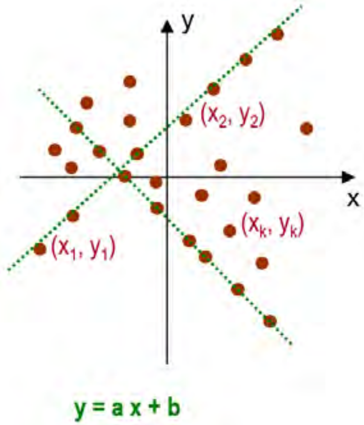
Vehicle Tracking and Identification

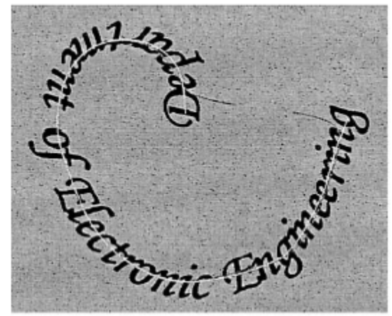


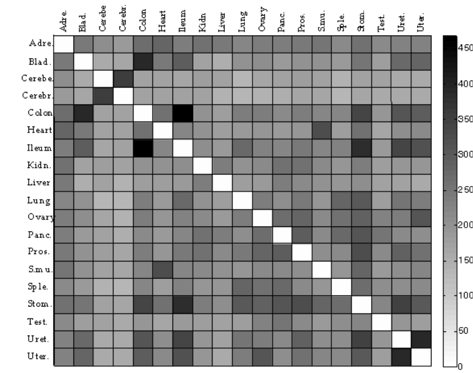
References: H Kung, KF Yang, L Chen, YJ Li, LLH Chan, and H Yan, *IEEE Trans Intelligent Transportation Systems*, 19(3):814-825, 2018.
H Kuang, C Liu, L Chen, LLH Chan, RCC Cheung, and H Yan, *IEEE Trans. SMC: Systems*, 49(1):71-80, 2019.

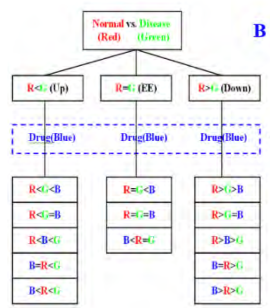
Co-Clustering

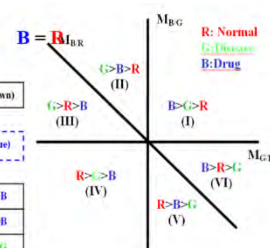




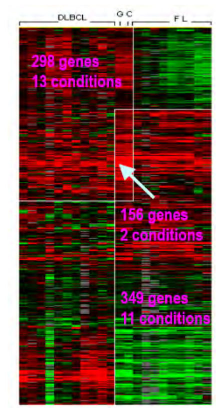




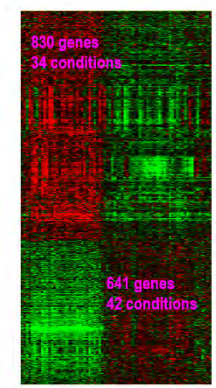




Slant axis: perfect drug
II, V: drug works, not enough
III, VI: drug works, too strong
I, IV: bad drug, side effect



Human Lymphoma Data



Breast Cancer Data

References:

H Yan, *US Patent* 7263538, 2007.

X Gan, A Liew and H Yan, *US Patent* 7849088, 2010.

X Gan, A Liew and H Yan, *Nucleic Acids Research*, 34:1608, 2006.

H Zhao and H Yan, *BMC Bioinformatics*, 8:256, 2007.

H Zhao, A Liew, X Xie and H Yan, *J Theoretical Biology*, 251:264-274, 2008.

D Wang and H Yan, *J Theoretical Biology*, 317:200, 2013.

W Yang, D Dai, and H Yan, *IEEE KDE*, 20:601, 2008.

W Yang, D Dai, and H Yan, *IEEE KDE*, 23:568, 2011.

P Tino, H Zhao, and H Yan, *IEEE/ACM TCBB*, 8:1093, 2011.

H Zhao, DD Wang, L Chen, X Liu, and H Yan, *PLoS ONE*, 11(9): e0162293, 2016.

H Yan, *IEEE Systems, Man and Cybernetics Magazine*, 3(2):23, 2017.

Our contributions:

* Biclutsering algorithms

* Co-expressed gene identification

* Cancer type sub-type classification

Cell Image Tracking and Lineage Analysis

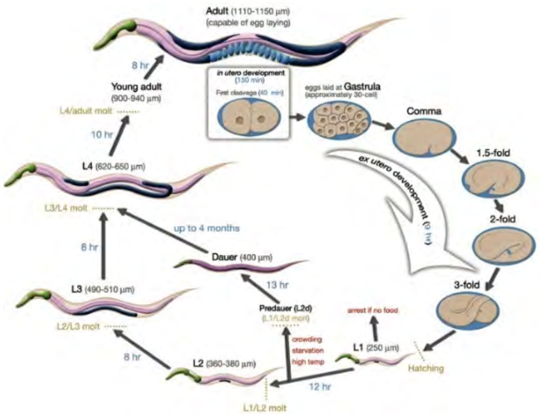
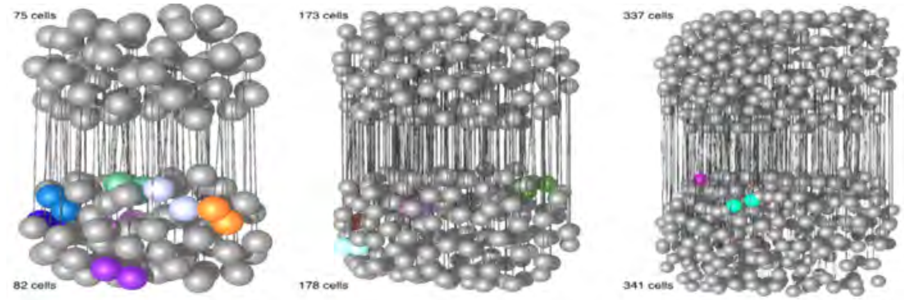
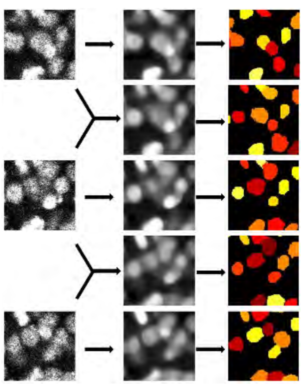
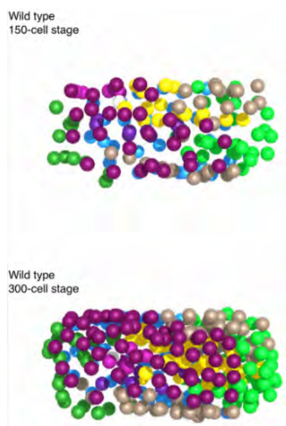


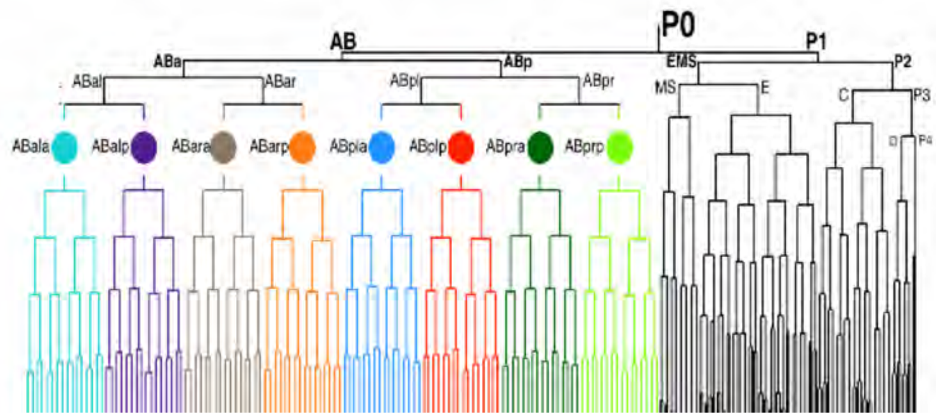
Image from www.wormatlas.org/ver1/handbook/anatomyintro/anatomyintro.htm

> 20 TB data









Collaborations:
HKBU
HKUST
CityU

Our contributions:
* Cell tracking and matching
* Cell-cell contact analysis
* Biological network inference

References:

VW S Ho, MK Wong, X An1, D Guan, J Shao, HCK Ng, X Ren, K He, J Liao, Y Ang, L Chen, X Huang, B Yan, Y Xia, LLH Chan, KL Chow, H Yan, and Z Zhao, *Molecular Systems Biology*,11:814, 2015.

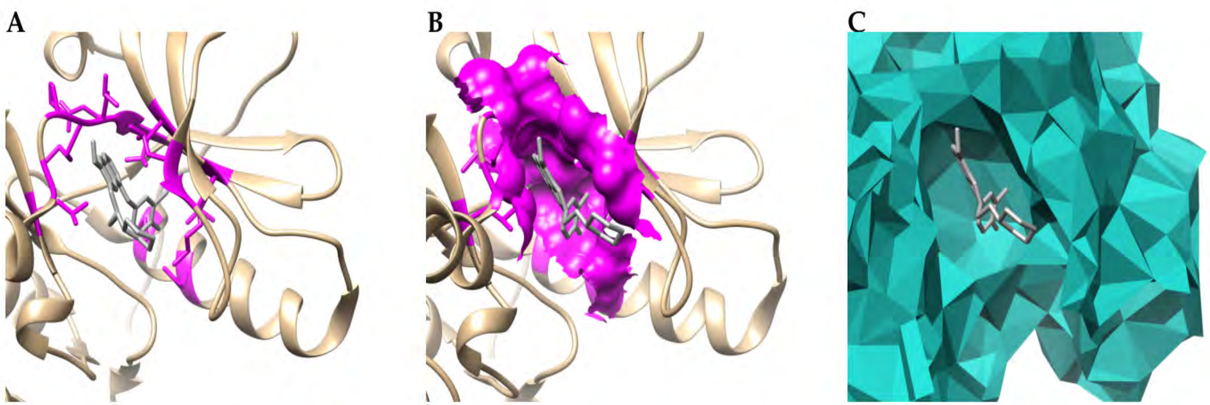
XT Huang, Y Zhu, LLH Chan, Z Zhao, and H Yan, *Molecular BioSystems*, 12:85-92, 2016.

L Chen, Z Zhao, and H Yan, *IEEE J. Selected Topics in Signal Processing*, 10(1):185-192, 2016.

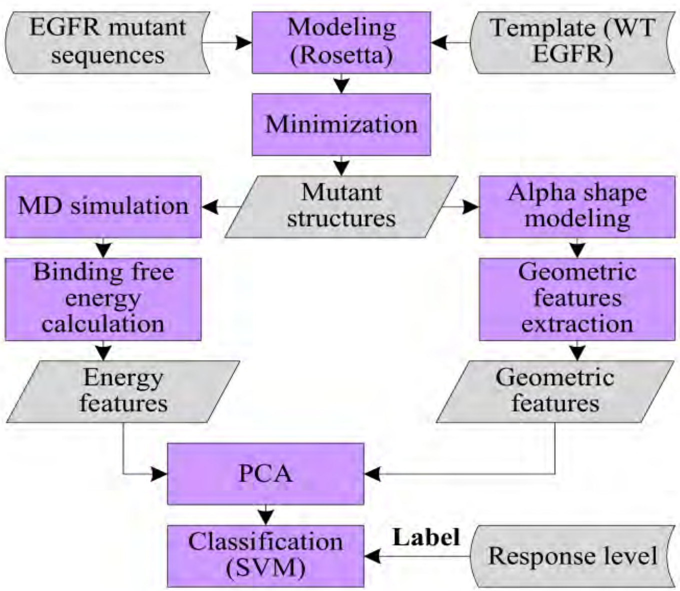
XT Huang, Y Zhu, LHL Chan, Z Zhao, and H Yan, *Bioinformatics*, 33(10):1528-1535, 2016.

L Chen, VWS Ho, MK Wong, X Huang, LY Chan, HCK Ng, X Ren, H Yan, and Z Zhao, *Genetics*, 209(1):36-49, 2019.

Cancer Drug Resistance



Feature	L858R (RL=2)	L858R_T790M(RL=3)
Curvature index 1	6	9
Curvature index 2	4	8
Curvature index 3	4	6
Curvature index 4	0.4700	0.5252
Curvature index 5	0.0876	0.0941
Curvature index 6	-0.6128	-0.5639
Curvature index 7	0.0754	0.0728
Mean connectivity	6.1148	6.4576
Connectivity variance	4.5934	4.2821
Atom index	0.5980	0.5728
VDWAALS	-51.6461	-47.1378
EEL	-26.1769	-9.8236
EGB	38.4677	29.7224
ESURF	-6.6548	-5.9291

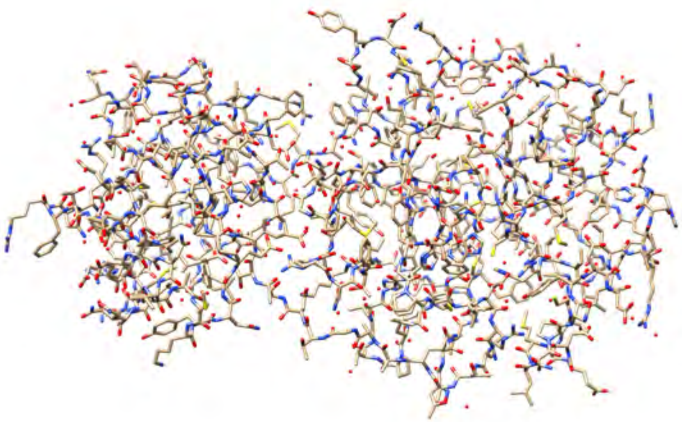


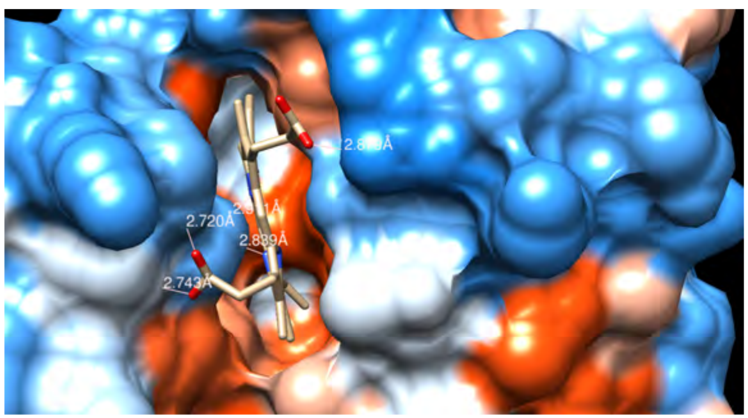
$$
\begin{aligned}
PC_1 = {} & 0.34 f_1 + 0.47 f_2 + 0.49 f_3 + 0.29 f_4 + 0.27 f_5 \\
& - 0.03 f_6 - 0.18 f_7 + 0.04 f_8 + 0.01 f_9 + 0.15 f_{10} \\
& + 0.08 f_{11} + 0.34 f_{12} - 0.28 f_{13} + 0.10 f_{14}
\end{aligned}
$$

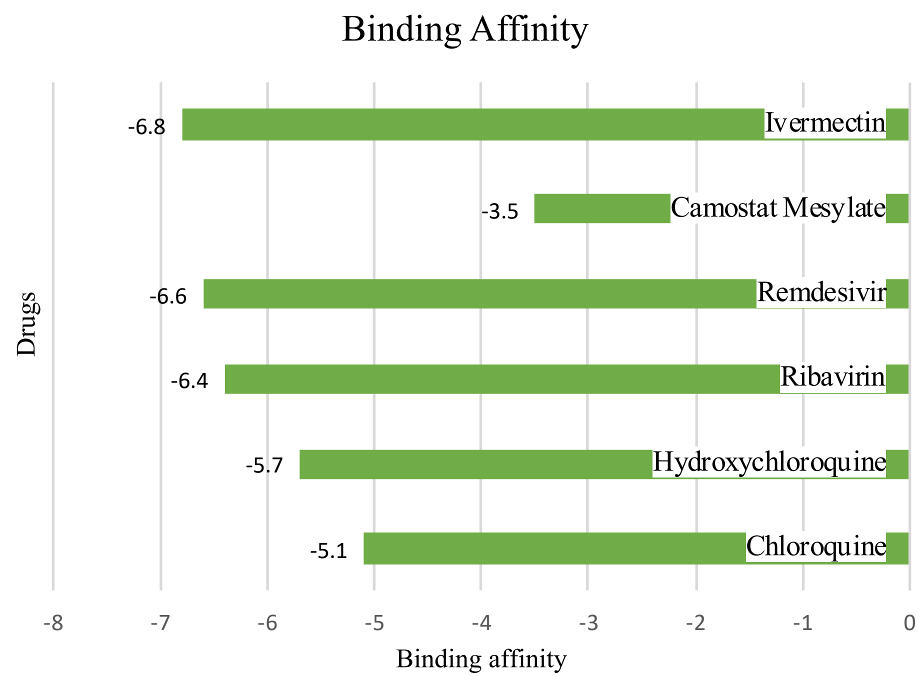
$$
\begin{aligned}
PC_2 = {} & 0.17 f_1 + 0.22 f_2 + 0.24 f_3 + 0.08 f_4 + 0.17 f_5 \\
& + 0.05 f_6 - 0.10 f_7 + 0.07 f_8 + 0.04 f_9 + 0.20 f_{10} \\
& - 0.25 f_{11} - 0.63 f_{12} + 0.48 f_{13} - 0.27 f_{14}
\end{aligned}
$$

Reference: L Ma, DD Wang, B Zou, and H Yan, *IEEE TCBB*, 14(5):1187-1194, 2017.

Research on COVID-19

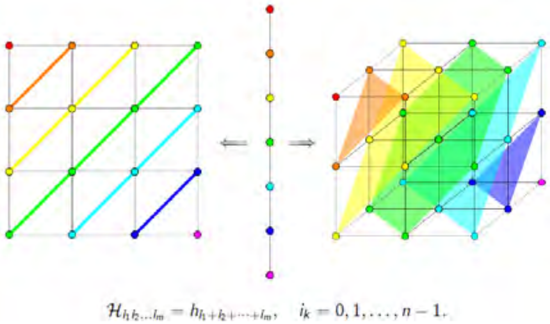


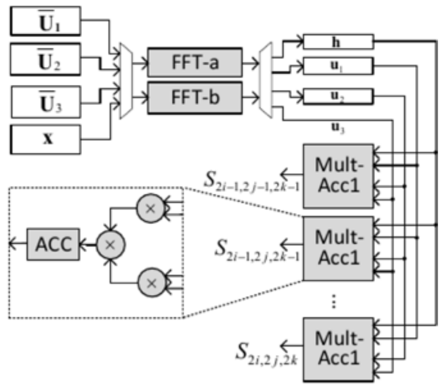


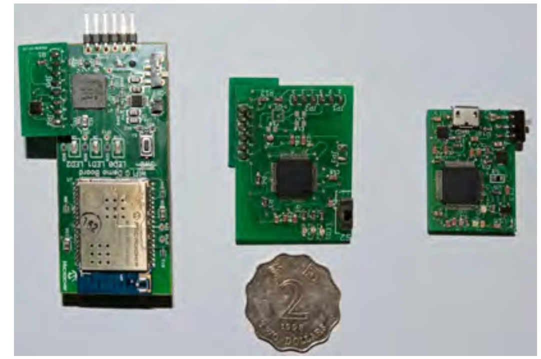


Reference: A Ghosh and H Yan, paper submitted.

Hardware Accelerators & Portable Electronic Device

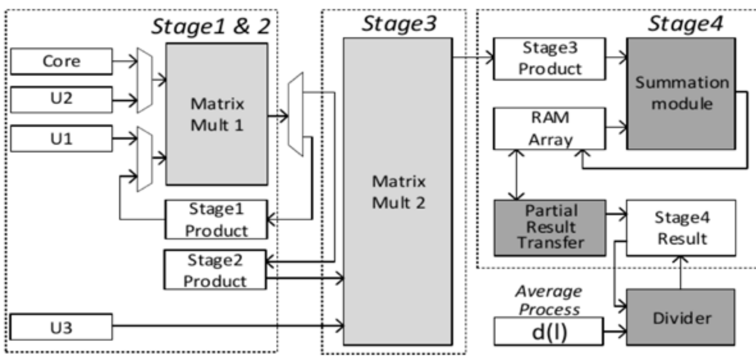


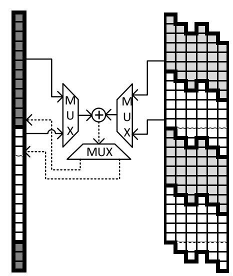




$$\hat{\mathcal{H}}(\mathbf{x}) = \mathcal{S} \times_1 \mathbf{U}_1^\top \times_2 \mathbf{U}_2^\top \times_3 \mathbf{U}_3^\top$$

$$\tilde{S}_{ijk} = ifft\{\tilde{\mathbf{x}}\}^\top \cdot [fft\{\bar{\mathbf{U}}_1(:,i)\}.*fft\{\bar{\mathbf{U}}_2(:,j)\}.*fft\{\bar{\mathbf{U}}_3(:,k)\}]$$





Storage change

$$N_1^2 + N_2^2 + N_3^2 \rightarrow N_1 + N_2 + N_3 - 2$$

Computational time change

$$O(N_1 r^3 + N_1 N_2 r^2 + N_1 N_2 N_3 r + N_1 N_2 N_3) \rightarrow O((k_3 + k_1 k_2 k_3 + 1)\frac{N_1 N_2 N_3 r_2}{k_1 k_2 k_3 PE_2})$$

References:

B Min, WP Huang, RCC Cheung, and H Yan, *Microelectronics Journal*, 85:25-33, 2019.

WP Huang, PY Kwan, W Ding, B Min, RCC Cheung, L Qi, and H Yan, *Microprocessors and Microsystems*, 64:120-127, 2019.

香港城市大學
City University of Hong Kong

Positions Available

- Matrix and tensor theories
- Data science and engineering
- Machine Learning, Pattern classification
- C++ and Python programming
- Efficient computational algorithms
- Hardware accelerators (GPU, FPGA, ASIC)
- Applications: image and biomedical data analysis

Contact: h.yan@cityu.edu.hk