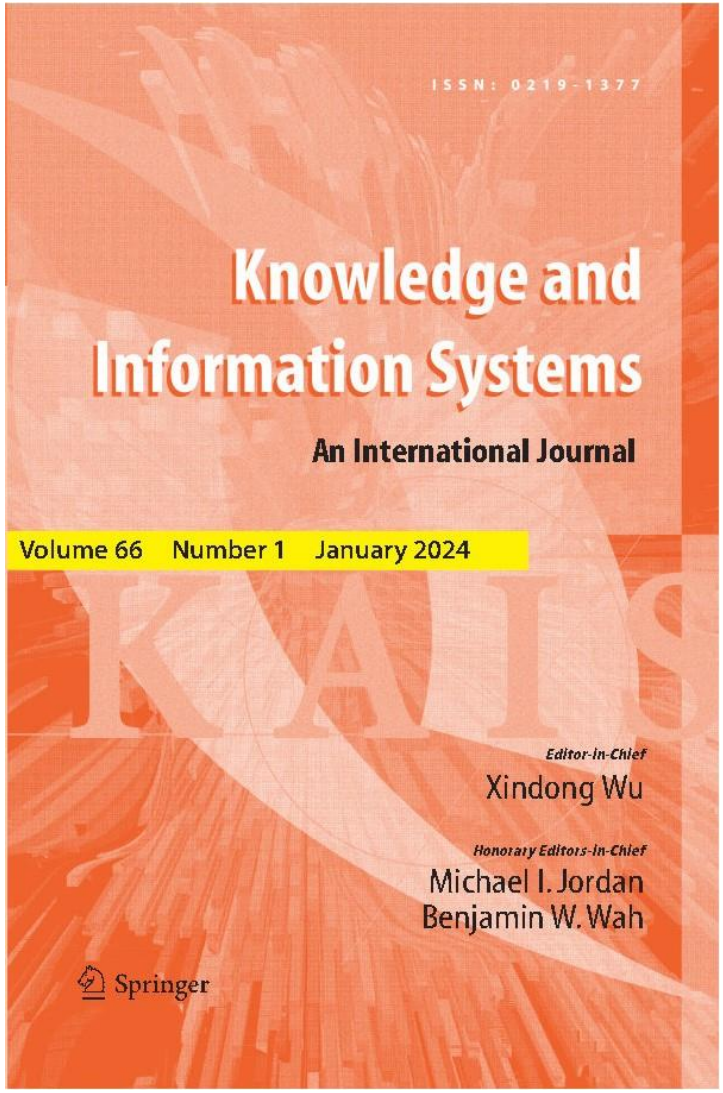
# ICDM '24 Award Ceremony

Steering Committee Chair:
Xindong Wu

1. ICDM Service Award
2. ICDM Research Award
3. Best Paper Awards (3)
4. 10-Year ICDM Impact Award
5. Tao Li Award
6. KAIS Journal Best Paper Award
7. ICDM Female Scholar Awards

# Major ICDM Activities in the Past 24 Years




① KAIS journal of the best papers every year since 2001 (the 2023 Impact Factor: 2.5; h-5: 48 by Google Scholar; both top 10!)

② IEEE ICDM Research Contributions Award and IEEE ICDM Outstanding Service Award since 2001(Nobel Prizes in Data Mining)

③ Data mining on ICDM submissions 2003 ~ 2015 (ICDM rules)

④ A panel discussion on "top 10 algorithms in data mining" in 2006 (Google citations: 7897 as of 12/07/24)

⑤ Workshops proceedings by IEEE CPS since 2006

⑥ Double blind reviewing since 2007 (and triple blind since 2011)

⑦ IEEE ICDM 10-Year Highest-Impact Paper Award since 2010

⑧ Tao Li Award since 2018

⑨ ICDM Female Scholar Awards since 2023

# 2024
# ICDM Steering Committee Awards

## (each with a check of $2,500)

- **IEEE ICDM Outstanding Service Award**
  - **highest recognition for service achievements in Data Mining.**
- **IEEE ICDM Research Contributions Award**
  - **the highest recognition for research achievements in Data Mining.**

# 2024 IEEE ICDM
## Nomination and Evaluation Committees

- Xindong Wu (Co-Chair), Hefei University of Technology, China
- Xingquan Zhu (Co-Chair), Florida Atlantic University, USA
- Charu Aggarwal, IBM T. J. Watson Research Center, New York, USA
- James Bailey, University of Melbourne, Australia
- Diane Cook, Washington State University, USA
- Peter Flach, University of Bristol, UK
- Eamonn Keogh, University of California Riverside, USA
- Jian Pei, Simon Fraser University, Canada
- Claudia Plant, University of Vienna, Austria
- Jie Tang, Tsinghua University, China
- Chengqi Zhang, University of Technology Sydney, Australia

# 2024 IEEE ICDM
# Outstanding Service Award

Professor Takashi Washio

Kansai University, Japan

- IEEE ICDM Steering Committee
- Conference Co-Chair for ICDM 2022
- Associate Editor, TKDE, TKDD, KAIS, DMKD

A talk by Takashi …


**IEEE ICDM-2024**

**Outstanding Service Award**

**December 11, 2024**

Takashi Washio

Kansai University, Japan

# Persistence pays off

# Service Overview along ICDM History

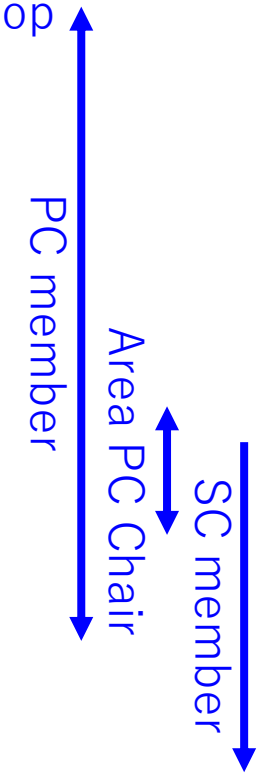
ICDM2001 (San Jose, USA) [The 1st ICDM]
ICDM2002 (Maebashi, Japan) : Tutorial chair, PC chair of a workshop

ICDM2006 (Hong Kong, China) : PC chair of a workshop

ICDM2013 (Dallas, USA) : Workshop co-chair
ICDM2014 (Shenzhen, China) : Organizer of a workshop
ICDM2015 (Atlantic City, USA) : Contest chair

ICDM2019 (Beijing, China) : Publicity chair

ICDM2022 (Orlando, USA) : General co-chair



# Some Major Services

- Tutorial chair @ ICDM2002 (Maebashi, Japan)
  - Organized 6 tutorials given by representative researchers.
  - Attracted many participants in the local organization to establish international reputation of ICDM.
- Workshop co-chair @ ICDM2013 (Dallas, USA)
  - Organized 19 workshops which cover cutting edge problems.
- Contest chair @ ICDM2015 (Atlantic City, USA)
  - Drawbridge cross-device connection challenge contest was held.
- Publicity chair @ ICDM2019 (Beijing, China)
  - Acquired a record-breaking number of 1046 paper submissions from 52 different countries/regions.
    194 papers were accepted (acceptance rate 18.5%).

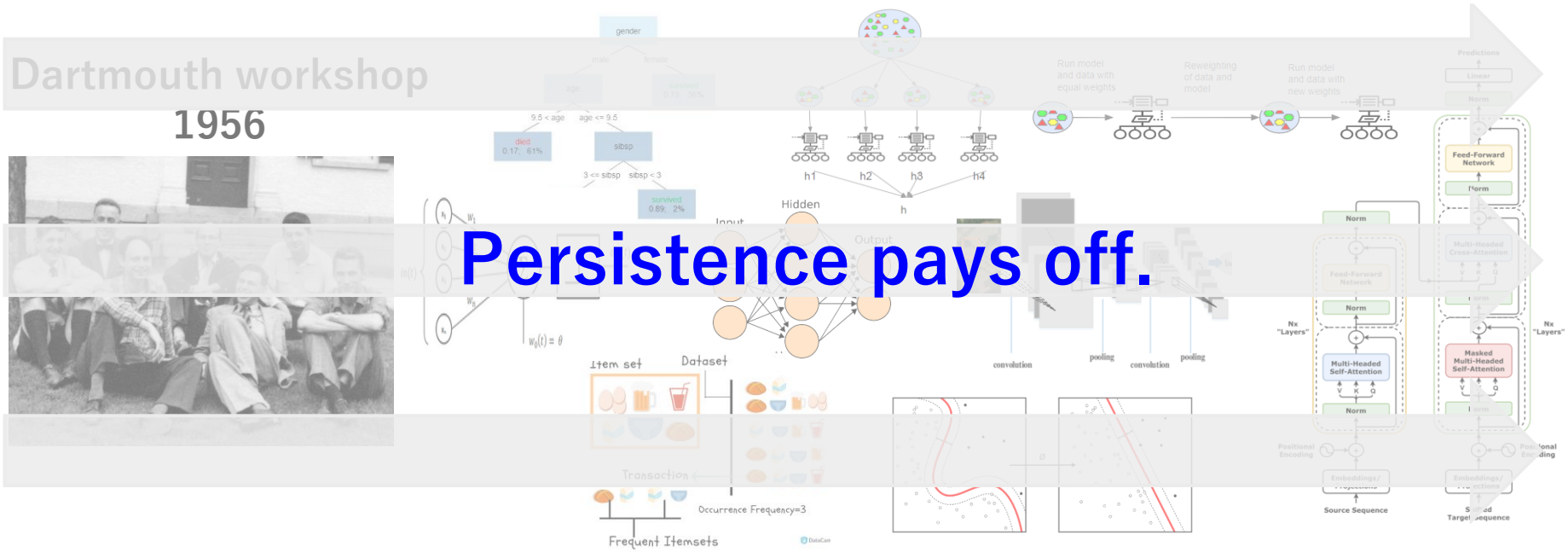
## Activities as an ICDM SC member, General co-chair of ICDM2022 and an Assoc. editor of major data mining journals

- Has been an ICDM steering committee member since 2016.
- Served as general co-chair of ICDM2022 (Orland, FL, USA)
  ICDM2022 was the first face-to-face meeting
  after the COVID-19 pandemic.
  Under strong support of My T. Thai (general co-chair),
  Xingquan Zhu and Sanjay Ranka (PC co-chairs),
  we could restore the conference to normal operations
  and further enhance the reputation of ICDM.


- Has been a senior steering committee member of PAKDD.
- Has been Associate Editor of major data mining journals including KAIS, TKDD, TKDE, and DMKD.

I express my utmost gratitude to the ICDM community.

I strongly hope that the community will continuously support ICDM and ICDM will last ever after.



# 2024 IEEE ICDM Research Contributions Award

Professor Geoff Webb, F'IEEE

Monash University, Australia

- MultiBoosting (*Machine Learning*, **40**(2000), 2: 159-196)
- OPUS search algorithm (*JAIR*, **3**(1995), 1: 431-465)
- IEEE ICDM 10-year Highest Impact Award (2023)

A talk by Geoff …


OMG!

# My journey

- Alternative high school
  - Forced to take computing
- Bachelor of Arts
  - Majoring in Philosophy
  - Sub major and Honours in Computer Science
- PhD on AI in Education
  - Focus on student modeling
- Realised I was developing machine learning algorithms
- Became a ML researcher
  - Primary interest was philosophical rather than practical
    - Superpower!
  - My first ML presentation was a trainwreck
    - Ross Quinlan ripped my experiments apart!
    - Learn from your mistakes!
  - Then ML became a BIG DEAL!

# Lessons

- Use your superpowers
- Learn from your mistakes
- Don’t follow the crowd
- Work with the best
- Luck plays a big role
- Life is what happens while you are busy making other plans

# Thanks!

- Nominators
- Awards & Steering Committees
- Students & Postdocs
- Collaborators
- Colleagues
- Mentors
- Friends
- Family
- Community

# ICDM '24 Awards Committee

- **Committee Chair: Geoff Webb** <geoff.webb@monash.edu>
  - 2023 Impact Paper Award Winner
- **Committee Members:**
  - Elena Baralis <elena.baralis@polito.it>
  - Diane Cook <cook@eecs.wsu.edu>
  - Vagelis Papalexakis <epapalex@cs.ucr.edu>
  - Hanghang Tong <htong@illinois.edu>
  - Xindong Wu <xwu@hfut.edu.cn>
  - Lina Yao <lina.yao@data61.csiro.au>
  - Kun Zhang <Kun.Zhang@mbzuai.ac.ae>

# ICDM '24 Best Paper Awards

- **ICDM '24 Best Paper Award ($1K from TCII):**
  - Scalable Graph Classification via Random Walk Fingerprints
  - by Peiyan Li, Honglian Wang, Christian Böhm
- **ICDM '24 Best Student Paper Award (2 * $500 from TCII):**
  - Normalizing Self-Supervised Learning for Provably Reliable Change Point DetectionP
  - by Alexandra Bazarova, Evgenia Romanenkova, Alexey Zaytsev
- **Best Student Paper Runner-Up (2 * $500 from TCII):**
  - Traffic Pattern Sharing for Federated Traffic Flow Prediction with Personalization
  - by Hang Zhou, Wentao Yu, Sheng Wan

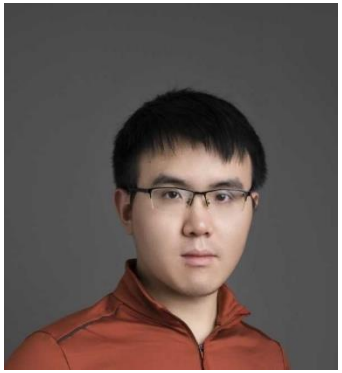
# 10 Year High Impact Award
## (with a check of $1,000)

- Selected by the Awards Committee
  - Considered some of the most cited papers that appeared in ICDM 2015.
  - A short list of papers were selected and ranked by members of the awards committee.
- The paper that stood out:
  - **A Unified Gradient Regularization Family for Adversarial Examples (ICDM-2015, pp. 301-309)**
  - ***By Chunchuan Lyu, Kaizhu Huang and Hai-Ning Liang***

A video talk by Chunchuan …

XJTLU Xi'an Jiaotong-Liverpool University 西交利物浦大学

# *A Unified Gradient Regularization Family for Adversarial Examples: After 9 Years*

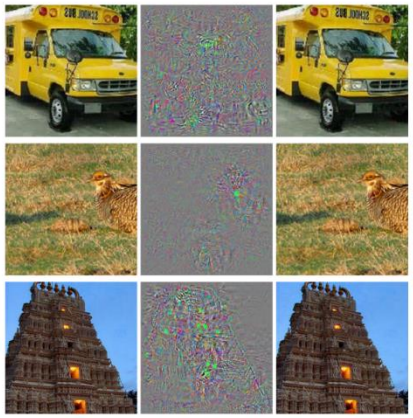


Chunchuan Lyu
Assistant Professor at
Xi'an Jiaotong Liverpool University
Chunchuan.Lyu@xjtlu.edu.cn

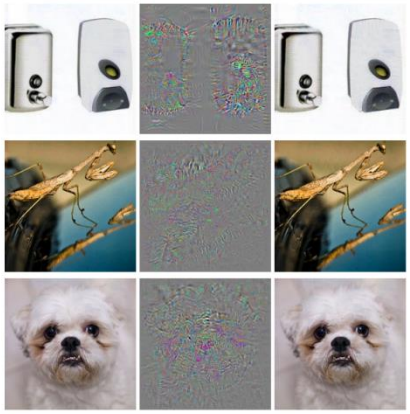
Hai-Ning Liang
Associative Professor at
HKUST (Guangzhou)
hainingliang@hkust-gz.edu.cn

Kaizhu Huang
Professor at
Duke Kunshan University
kaizhu.huang@dukekunshan.edu.cn

# Adversarial Examples



(a)



(b)

- Minimize $\|r\|_2$ subject to:
  1. $f(x+r) = l$
  2. $x + r \in [0,1]^m$

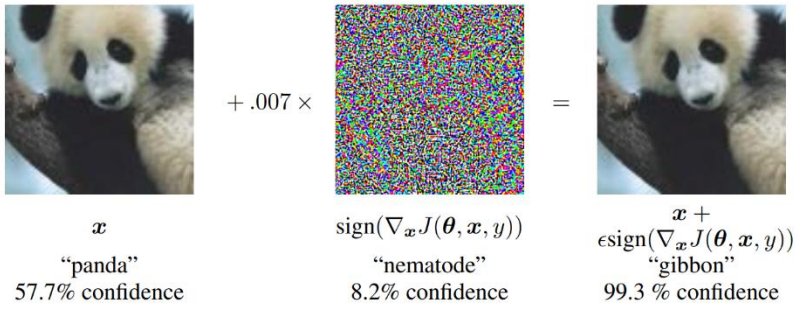
Why they exist?

How to defend?

Why vision works at all?

Intriguing properties of neural networks
Christian Szegedy et.al. ICLR 2014

# Adversarial Examples




EXPLAINING AND HARNESSING ADVERSARIAL EXAMPLES
Ian Goodfellow et.al. ICLR 2015

Very efficient adversarial examples finder.

Explain the adversarial examples by the linear behavior of approximating function.

# Minmax View

## Gradient Regularization Family

### Classification problem

$$\min_{\theta} \mathcal{L}(\boldsymbol{x} + \epsilon; \theta)$$

### Main framework for robutstness

$$\min_{\theta} \max_{\epsilon:\|\epsilon\|_p \leq \sigma} \mathcal{L}(\boldsymbol{x} + \epsilon; \theta)$$

- $\mathcal{L}(\boldsymbol{x}; \theta)$ [1] : loss function
- $\boldsymbol{x}$: data
- $\theta$: model parameters
- $\epsilon$: small perturbation

Ideally we would like to solve the min-max problem , if we try to build a robust model against any $\epsilon$.

[1] This notation might be shorthanded as $\mathcal{L}(\boldsymbol{x})$ or $\mathcal{L}$ in the following text.



Our paper was one of the first to propose this simplified minmax view.

In retrospect, this is the dual formulation of the original formulation.

# Minmax View

## Gradient Regularization Family

**Linearized main framework**

$$\min_{\theta} \max_{\epsilon:\|\epsilon\|_p \leq \sigma} \mathcal{L}(\boldsymbol{x}) + \nabla_{\boldsymbol{x}} \mathcal{L}^T \epsilon$$

**Solution of approximated max problem**

$$\epsilon = \sigma \ \text{sign}(\nabla \mathcal{L})(\frac{|\nabla \mathcal{L}|}{\|\nabla \mathcal{L}\|_{p^*}})^{\frac{1}{p-1}}$$

**Putting perturbation back**

$$\min_{\theta} \mathcal{L}(\boldsymbol{x}) + \sigma \|\nabla \mathcal{L}\|_{p^*}$$

Instead of minimizing $\mathcal{L}(\boldsymbol{x})$, we try to minimize $\mathcal{L}(\boldsymbol{x} + \epsilon)$ where $\epsilon$ is defined above.

Historical notes:

We saw the Intriguing Paper in my junior year 2013.

Spent a whole year thinking about this problem.

Came to the l2 perturbation by intuition.

Saw the panda from Ian on a very dark cold winter day in 2014.

Realized that well, we have to reformulate it in a more general framework.

Got rejected by UAI due to my terrible writing (reviewer: just another regularizer)

Missed ICDM banquet twice

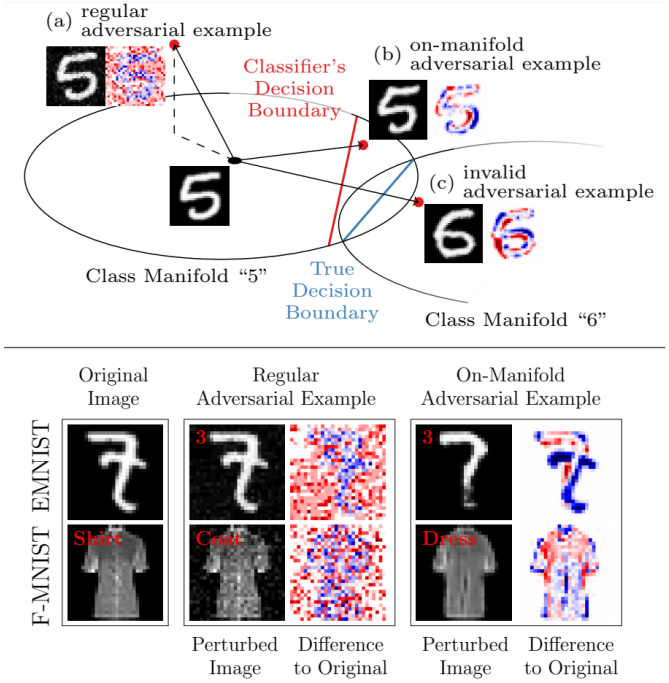
## What happens afterwards?

Not inclusive, leaving out best papers and papers from Prof. Kaizhu Huang who continues to work on adversarial examples afterwards.

I restarted on NLP (the "interesting" parts killed/solved by LLMs), now restarting again on Reinforcement Learning.

So, I am not an expert anymore, but I take this chance to raise a question.

# Manifold View




Disentangling Adversarial Robustness and Generalization
David Stutz, et al. CVPR 2019

1. on-manifold adversarial examples, exist;
2. on–manifold adversarial examples are generalization errors, and on-manifold adversarial training boosts generalization;

Experiments are limited, but the intuition is reasonable

# Pull Adversarial Examples to the Manifold

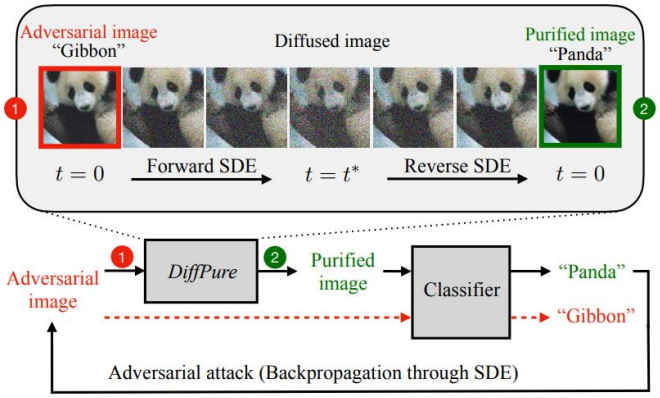
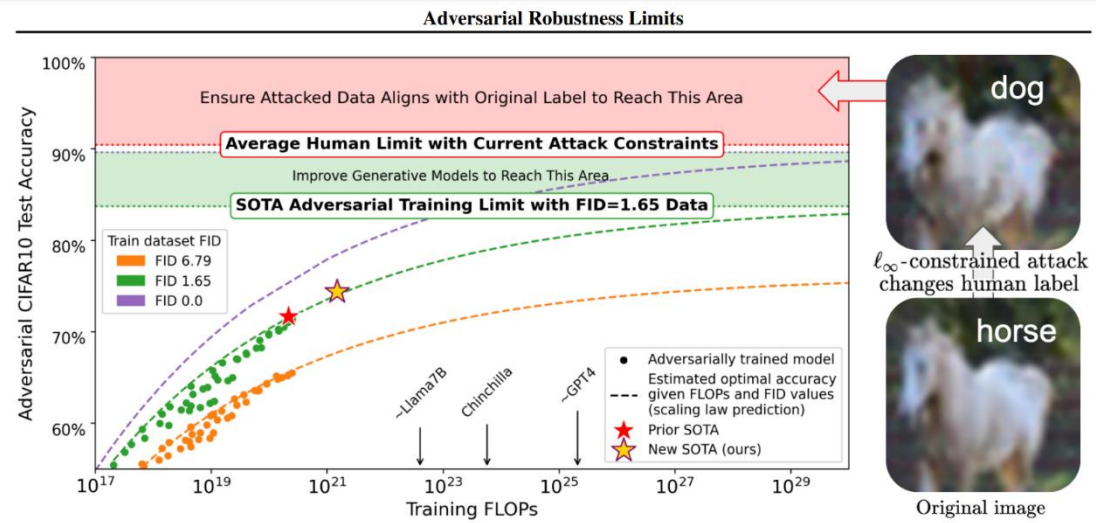



*Figure 1.* An illustration of *DiffPure*. Given a pre-trained diffusion model, we add noise to adversarial images following the forward diffusion process with a small diffusion timestep $t^*$ to get diffused images, from which we recover clean images through the reverse denoising process before classification. Adaptive attacks backpropagate through the SDE to get full gradients of our defense system.

Diffusion Models for Adversarial Purification
Weili Nie et al
ICML 2022

Maybe, we are not so susceptible to adversarial examples because **our brains are constantly denoising/purifying our data.**

# Scaling Law




Adversarial Robustness Limits via Scaling-Law and Human-Alignment Studies
Brian R. Bartoldson et al
ICML 2024

Used towns of data augmentation.

Did not use purification.

To people with tons of GPUs:

Could you please study the scaling law for purification based model against adversarial examples?

We can not only scale the training time compute, but also the testing time purification steps.

If the scaling law suggests that we can close the gap between standard accuracy and robust accuracy, we would have understood the problem in principle.

We would know that purification pulled off-manifold data back, and on-manifold data are simply about generalization.

**Chunchuan Lyu**
**chunchuan.lyu@xjtlu.edu.cn**

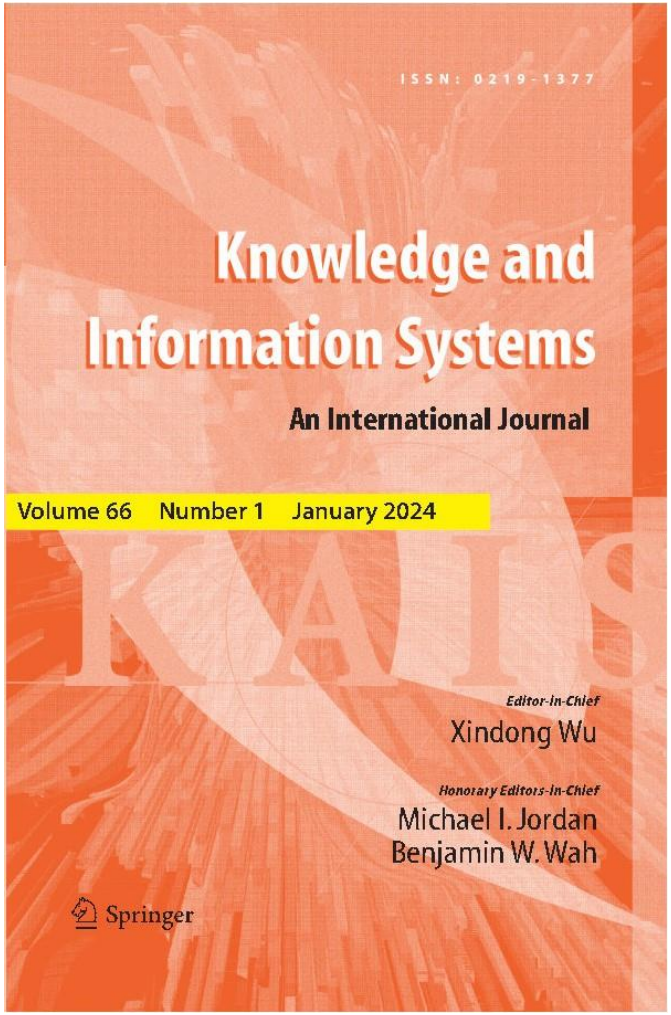
# Thank you!

# 2024 KAIS Journal Best Paper Award

## (A check of $1K from Springer)

Best Paper published in 2021:

- Model Complexity of Deep Learning: A Survey
- by Xia Hu, Lingyang Chu, Jian Pei, Weiqing Liu and Jiang Bian
- *Knowledge and Information Systems*, Volume 63, Issue 10, Pages 2585-2619



# 2024 Tao Li Award
## (a $2K check from Tao Li Foundation)

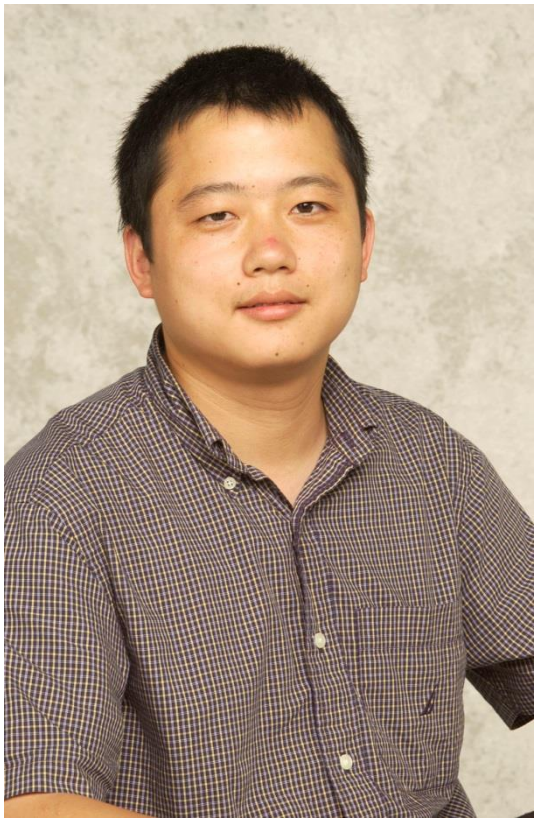
Was selected by a committee

- Tao Li  Award Committee Chair: Xingquan Zhu

2024 Tao Li Award winner:

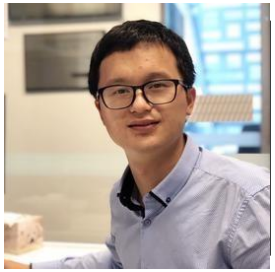
- Shirui Pan

A short talk by Shirui …



Griffith UNIVERSITY
Queensland, Australia

# Advancing AI and Data Mining with Graph Neural Networks



**Shirui Pan**

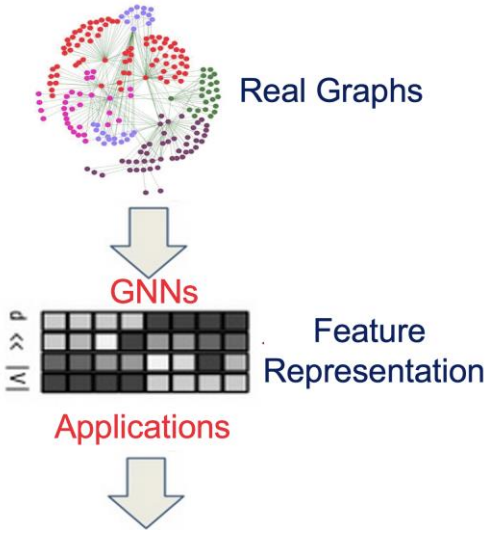
Griffith University, Australia

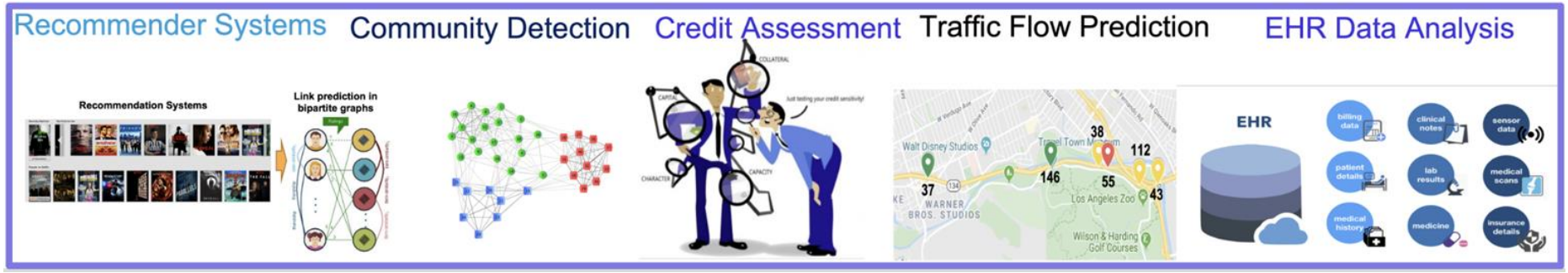
https://shiruipan.github.io/

(Generated by DALL·E)

# Graph Neural Networks

- Methods and Applications
  - Core Deep Learning Technique
  - Effective Representation for Graph Data
  - Wide applications



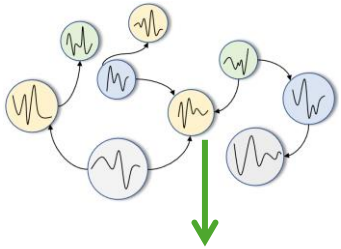



# Graph Neural Networks for Spatial Temporal Analysis

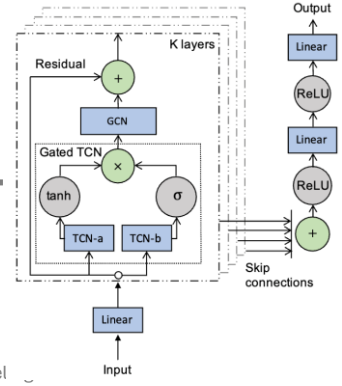
- **Graph WaveNet (IJCAI-2019)**

1. Explicitly Capturing the Spatial and Temporal Dependence
2. GNN-based Model for Traffic Forecasting

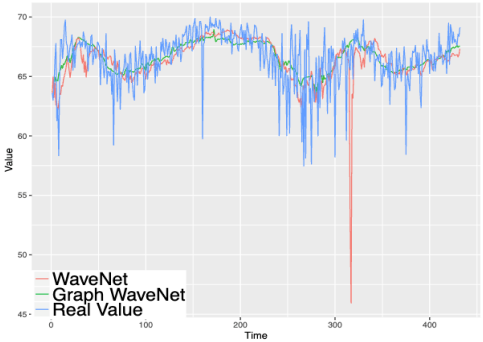
Graph Modeling



Graph WaveNet Training




Prediction




- 2,200+ Citations
- The *Most Cited* IJCAI-2019 Paper
- Still outperforms many recent models in a recent evaluation (TKDE-24)

| | Source | PEMS04 | | | PEMS08 | | |
|---|---|---|---|---|---|---|---|
| | | MAE | RMSE | MAPE | MAE | RMSE | MAPE |
| GWNet | [22], [62], [23], [63], [57] | 25.45 | 39.70 | 17.29% | 19.13 | 31.05 | 12.68% |
| | [25], [64] | 24.89 | 39.66 | 17.29% | 18.28 | 30.04 | 12.15% |
| | [46] | 19.36 | 31.72 | 13.31% | 15.07 | 23.85 | 9.51% |
| | [24] | 28.15 | 39.88 | 18.52% | 20.30 | 30.82 | 13.84% |
| | BasicTS+ | **18.80** | **30.14** | **13.19%** | **14.67** | **23.55** | **9.46%** |
| | Gap | **33.21%↑** | **24.42%↑** | **28.78%↑** | **27.73%↑** | **23.59%↑** | **31.64%↑** |

Results from Shao et al (TKDE 2024)

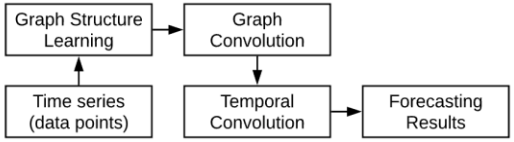


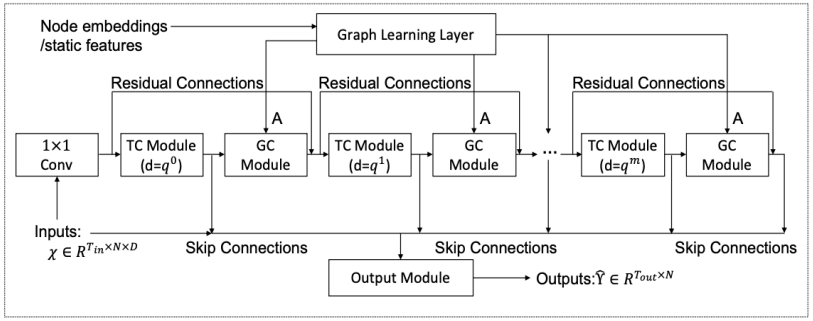
Wu, Z., Pan, S., Long, G., Jiang, J., & Zhang, C. (2019). Graph wavenet for deep spatial-temporal graph model

# Graph Neural Networks for Time Series

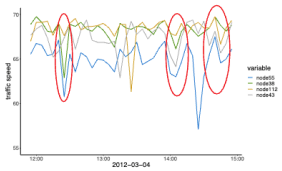
- **MTGNN (KDD-2020)**

1. First GNN model for multivariate time series data
2. Automatically learn the graph structure from the data
3. Can be used in any multivariate time series

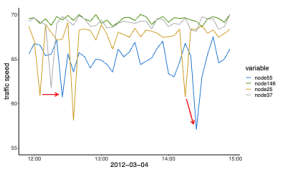


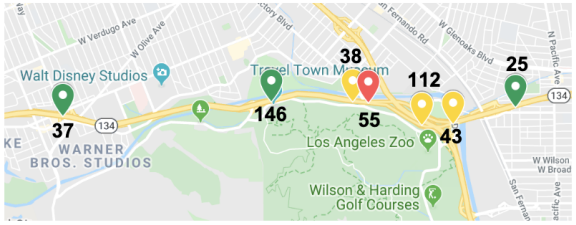

- 1,500+ Citations
- The *Most Cited* KDD 2020 Paper



(a) Time series of node 55 and its top-3 neighbors given by the pre-defined A.



(b) Time series of node 55 and its top-3 neighbors given by the learned A.




(c) Node locations of node 55 and its neighbors marked on Google Maps. Yellow nodes represent node 55's top-3 neighbors given by the pre-defined A. Green nodes represent node 55's top-3 neighbors given by the learned A.

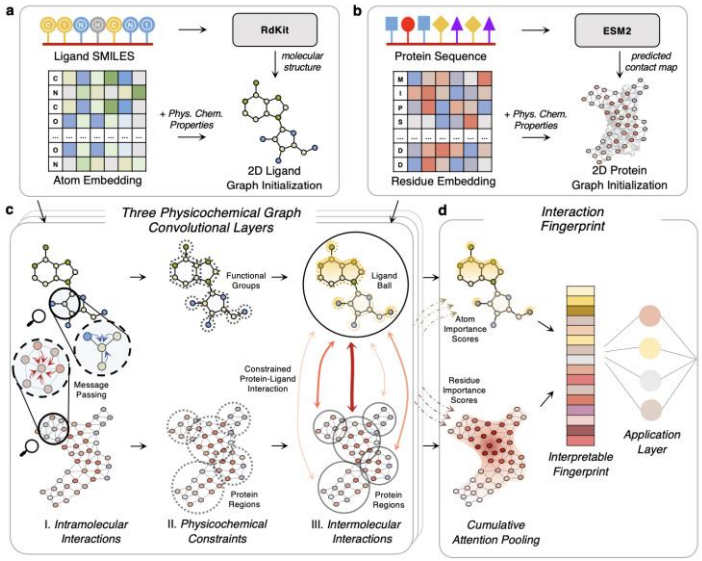


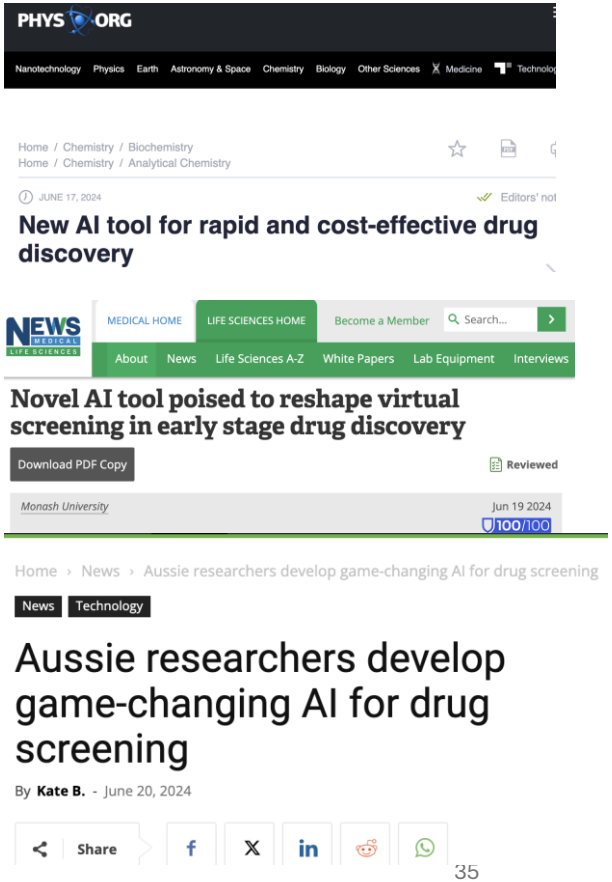
Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020). Connecting the dots: Multivariate time series forecasting with graph neural networks. In KDD.

# Graph Neural Networks for Drug Discovery

- **PSICHIC (Nature Machine Intelligence - 2024)**

1. Predicts the affinity between molecules and proteins
2. Enables fast and accurate virtual screening






Koh, H.Y., Nguyen, A.T.N.#, Pan, S.#, May, L.T.#, Webb, G.I.#, 2024. Physicochemical graph neural network for learning protein–ligand interaction fingerprints from sequence data. Nature Machine Intelligence.

# Beyond Technical Contributions

IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, VOL. 32, NO. 1, JANUARY 2021

## A Comprehensive Survey on Graph Neural Networks

Zonghan Wu, Shirui Pan, *Member, IEEE*, Fengwen Chen, Guodong Long, Chengqi Zhang, *Senior Member, IEEE*, and Philip S. Yu, *Life Fellow, IEEE*

TNNLS 2021
10,000 Citations
2024IEEE TNNLS Outstanding Paper Award

## A Survey on Knowledge Graphs: Representation, Acquisition and Applications

Shaoxiong Ji, Shirui Pan, *Member, IEEE,* Erik Cambria, *Senior Member, IEEE,* Pekka Marttinen, Philip S. Yu, *Life Fellow, IEEE*

- TNNLS 2022
- 2,400 Citations

3580 IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 36, NO. 7, JULY 2024

## Unifying Large Language Models and Knowledge Graphs: A Roadmap

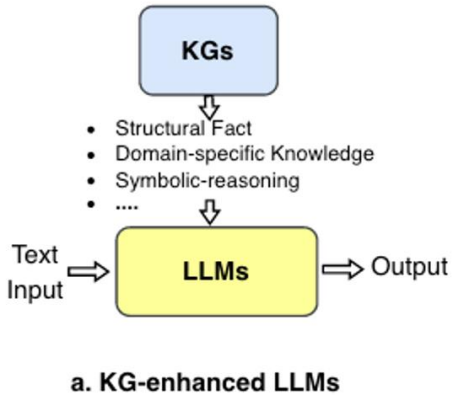
Shirui Pan, *Senior Member, IEEE*, Linhao Luo, Yufei Wang, Chen Chen, Jiapu Wang, and Xindong Wu, *Fellow, IEEE*

- TKDE 2024
- 600 Citations

# Thank You

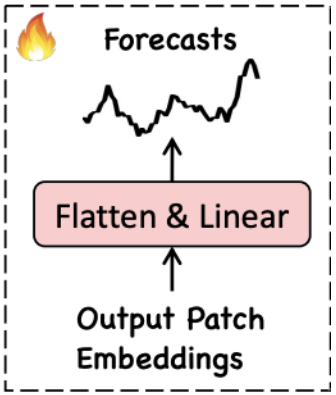
## Future Work

- **Graphs/Time Series in the Era of Large Language Models**

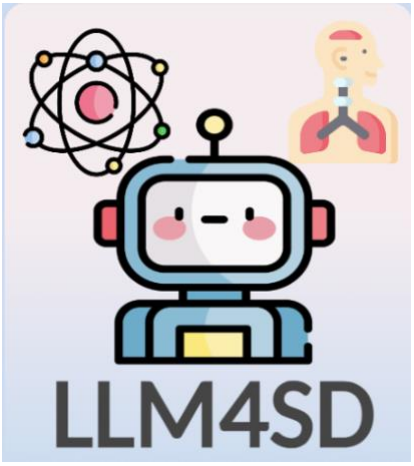



a. KG-enhanced LLMs

(KNOWLEDGE) GRAPHS + LLM




TIME SERIES FOUNDATION MODELS




GNN/LLM + SCIENCE


# ICDM Female Student Awards in Data Mining

- Selected by the Award Committee
  - Xiaofeng Gao, Co-Chair <gao-xf@cs.sjtu.edu.cn>
  - Xiaochun Yang, Co-Chair <yangxc@mail.neu.edu.cn>
  - Elena Baralis <elena.baralis@polito.it>
  - Xiuzheng Cheng <xzcheng@sdu.edu.cn>
  - My Thai <mythai@cise.ufl.edu>
  - Bhavani Thuraisingham <bxt043000@utdallas.edu>

# 2024 Young Female Student Award (a check of $1K)

**Li Qian** (Ph.D. Candidate, Ludwig-Maximilians-Universität München, Germany)

- **Remarkable Performance at ICDM**
  - **Li Qian**, et al., DynoGraph: Dynamic Graph Construction for Nonlinear Dimensionality Reduction, **ICDM'24**
  - **Li Qian**, et al., ADOD: Adaptive Density Outlier Detection, **ICDM'24**
  - **Li Qian**, et al., Density-based Clustering for Adaptive Density Variation, **ICDM'21**
- **Extensive Experience across Academia and Industry**
  - Engineer Roles:
    - Big Data Architect Assistant & Network Engineer Assistant
  - Academic Roles**:**
    - Junior Member at Munich Center for Machine Learning
    - Visiting Researcher at the University of Vienna


# Thanks!

## IEEE International Conference on Data Mining

The world's premier research conference in data mining