

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY) 04-11-2022		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 25-Jan-2021 - 24-Jan-2022	
4. TITLE AND SUBTITLE Final Report: A GPU/CPU Hybrid Cluster for Accelerating Network-of-X Research and Education			5a. CONTRACT NUMBER W911NF-21-1-0088		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER 611103		
6. AUTHORS			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES University of Illinois - Urbana - Champaign c/o Office of Sponsored Programs 1901 S. First Street, Suite A Champaign, IL 61820 -7406			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 77318-HC-RIP.14		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:		17. LIMITATION OF ABSTRACT		15. NUMBER OF PAGES	
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU	UU	19a. NAME OF RESPONSIBLE PERSON Hanghang Tong	
				19b. TELEPHONE NUMBER 217-244-0884	

RPPR Final Report

as of 04-Nov-2022

Agency Code: 21XD

Proposal Number: 77318HCRIP

Agreement Number: W911NF-21-1-0088

INVESTIGATOR(S):

Name: Hanghang Tong
Email: htong@illinois.edu
Phone Number: 2172440884
Principal: Y

Organization: **University of Illinois - Urbana - Champaign (UIUC)**

Address: c/o Office of Sponsored Programs, Champaign, IL 618207406

Country: USA

DUNS Number: 041544081

EIN: 376000511

Report Date: 24-Apr-2022

Date Received: 04-Nov-2022

Final Report for Period Beginning 25-Jan-2021 and Ending 24-Jan-2022

Title: A GPU/CPU Hybrid Cluster for Accelerating Network-of-X Research and Education

Begin Performance Period: 25-Jan-2021

End Performance Period: 24-Jan-2022

Report Term: 0-Other

Submitted By: Hanghang Tong

Email: htong@illinois.edu

Phone: (217) 244-0884

Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees: 3

STEM Participants: 35

Major Goals: In the age of AI and big data, networks provide powerful context to connect different things (e.g., entities, datasets, models) together. This phenomenon is referred to as Network-of-X in this project. By linking different entities, datasets, and data mining models together, Network-of-X brings a unique opportunity to dramatically enrich the modeling power, boost the mining performance, and broaden the reasoning capabilities.

The project aims to build a networked infrastructure to accelerate Network-of-X research and education. The proposed infrastructure has three mutually complementary modules. The first module (Distributed Graph Database) consists of a storage-optimized cluster offering high disk throughput and a distributed graph database management system supporting fast transaction processing. The second module (Graph Computation Engine) consists of a GPU server that is computationally optimized. It supports two types of fundamental computations in Network-of-X, including the large sparse matrix computations and the deep learning computations. The third module (Interactive Graph Analytics) develops an integrated human-computer interactive platform that allows the end-users to effectively collaborate and coordinate with computers in decision making, cognitive modeling, and algorithm training and diagnosis.

Accomplishments: Using the built infrastructure from this project, including four dell PowerEdge C4140 GPU servers and 1 Dell 6.4TB NVMe Express Flash, we have advanced the research in network-of-X on several topics, including active graph neural networks, graph structure learning, graph contrastive learning, fair graph learning, knowledge graph reasoning, uncertain quantification on graphs, bundle recommendation, cross-layer dependence inference, equation learning, and few-shot learning on graphs. In total, it has led to 12 publications on these topics. These research outcomes are parts of the key deliverables in two DoD projects, including disinformation detection (DHS 17STQAC00001-03-03) and influence campaign and sense-making (DARPA INCAS HR001121C0165). In addition, the built infrastructure has benefited the following federal-sponsored projects, including network robustification (NSF 1947135), fair network learning (NSF 1939725), optimal deep graph learning (NSF 2134079), and AI for agriculture (NIFA 2020-67021-32799).

RPPR Final Report as of 04-Nov-2022

Training Opportunities: In terms of education and training, we have used the built infrastructure as a development and testing platform for training 25 graduate students and 10 undergraduate students in various topics of large-scale graph mining and learning. During the project period, we graduated two Ph.D students (Dr. Si Zhang and Dr. Boxin Du) and one master's student (Zhe Xu). In addition, research opportunities were provided to both graduate and undergraduate students in the form of course projects in CS512: Data Mining Principles that PI taught in spring 2021. We also host a visiting student (Haoran Li) from Arizona State University. PI and his students have given a tutorial on the topic of algorithmic fairness on network learning at ACM KDD 2022 and another tutorial on new frontier of multi-network mining at ACM KDD 2021. PI gave a plenary talk at Artificial Intelligence, Machine Learning and Data Science World Forum 2022 on the topic of "NetFair: Toward the Why Question of Network Mining". He gave a keynote talk at Meta on the topic of "User Response Prediction: A Graph-based Approach" and another keynote talk at the Deep Learning on Graphs: Method and Applications (DLG-AAAI'21) on the topic of "NetFair: Towards Fair Network Mining".

Results Dissemination: Dissemination

- In total, this project has led to 12 publications, which have been or will be presented at the top conferences in data mining and AI (e.g., KDD, WWW, AAAI, NeuIPS).
- PI and his students have given a tutorial on the topic of algorithmic fairness on network learning at ACM KDD 2022 and another tutorial on new frontier of multi-network mining at ACM KDD 2021.
- PI gave a plenary talk at Artificial Intelligence, Machine Learning and Data Science World Forum 2022 on the topic of "NetFair: Toward the Why Question of Network Mining".
- PI gave a keynote talk at Meta on the topic of "User Response Prediction: A Graph-based Approach" and another keynote talk at the Deep Learning on Graphs: Method and Applications (DLG-AAAI'21) on the topic of "NetFair: Towards Fair Network Mining".
- PI served as a panelist to discuss challenges and opportunities of the fairness in outlier detection at ODD workshop at KDD 2021.
- PI served as ACM CIKM2021 program committee co-chair.
- PI co-organized AI and social responsibility symposium.
- PI co-organized a workshop on Graph Techniques for Adversarial Activity Analytics at IEEE BigData 2021.

Honors and Awards: •PI received Dean's Award for Excellence in Research.

- PI was elevated to IEEE Fellow, for "contributions to graph mining".
- PI received Excellent Paper Award from IEEE Big Data Mining and Analytics.
- PI received Best Student Paper Runner-up from IEEE ICDM 2021.
- PI received Visa Research Faculty Award.
- Master student Yuheng Zhang received "Kuck Outstanding Master's Thesis Award".
- Ph.D student Jian Kang was selected as a Mavis Future Faculty Fellow and was recognized as UChicago Rising Stars in Data Science.
- Master student Shengyu Feng was recognized as a Siebel scholar and received UIUC Graduate College Conference Presentation Award.
- Ph.D student Yuheng Zhang received UIUC Graduate College Conference Presentation Award.

Protocol Activity Status:

RPPR Final Report
as of 04-Nov-2022

Technology Transfer: Nothing to Report

PARTICIPANTS:

Participant Type: PD/PI

Participant: Hanghang Tong

Person Months Worked: 1.00

Funding Support:

Project Contribution:

National Academy Member: N

Participant Type: Graduate Student (research assistant)

Participant: Zhe Xu

Person Months Worked: 6.00

Funding Support:

Project Contribution:

National Academy Member: N

Participant Type: Graduate Student (research assistant)

Participant: Yuheng Zhang

Person Months Worked: 3.00

Funding Support:

Project Contribution:

National Academy Member: N

Participant Type: Graduate Student (research assistant)

Participant: Shengyu Feng

Person Months Worked: 3.00

Funding Support:

Project Contribution:

National Academy Member: N

Participant Type: Graduate Student (research assistant)

Participant: Jian Kang

Person Months Worked: 6.00

Funding Support:

Project Contribution:

National Academy Member: N

Participant Type: Graduate Student (research assistant)

Participant: Lihui Liu

Person Months Worked: 8.00

Funding Support:

Project Contribution:

National Academy Member: N

Participant Type: Graduate Student (research assistant)

Participant: Qinghai Zhou

Person Months Worked: 2.00

Funding Support:

Project Contribution:

National Academy Member: N

RPPR Final Report
as of 04-Nov-2022

Participant Type: Research Experience for Undergraduates (REU) Participant

Participant: Zhenning Zhang

Person Months Worked: 5.00

Funding Support:

Project Contribution:

National Academy Member: N

Participant Type: Graduate Student (research assistant)

Participant: Boxin Du

Person Months Worked: 2.00

Funding Support:

Project Contribution:

National Academy Member: N

Participant Type: Graduate Student (research assistant)

Participant: Yuchen Yan

Person Months Worked: 9.00

Funding Support:

Project Contribution:

National Academy Member: N

Participant Type: Graduate Student (research assistant)

Participant: Yikun Ban

Person Months Worked: 1.00

Funding Support:

Project Contribution:

National Academy Member: N

Participant Type: Other Professional

Participant: Haoran Li

Person Months Worked: 6.00

Funding Support:

Project Contribution:

National Academy Member: N

ARTICLES:

RPPR Final Report as of 04-Nov-2022

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: WWW '22: Proceedings of the ACM Web Conference 2022
Publication Identifier Type: **Publication Identifier:**
Volume: **Issue:** **First Page #:**
Date Submitted: 11/4/22 12:00AM **Date Published:** 5/1/22 5:00AM
Publication Location:

Article Title: Graph Sanitation with Application to Node Classification

Authors: Zhe Xu, Boxin Du, Hanghang Tong

Keywords: graph learning, graph neural network, optimization

Abstract: The past decades have witnessed the prosperity of graph mining, with a multitude of sophisticated models and algorithms designed for various mining tasks, such as ranking, classification, clustering and anomaly detection. Generally speaking, the vast majority of the existing works aim to answer the following question, that is, given a graph, what is the best way to mine it? In this paper, we introduce the graph sanitation problem, to answer an orthogonal question. That is, given a mining task and an initial graph, what is the best way to improve the initially provided graph? By learning a better graph as part of the input of the mining model, it is expected to benefit graph mining in a variety of settings, ranging from denoising, imputation to defense. We formulate the graph sanitation problem as a bilevel optimization problem, and further instantiate it by semi-supervised node classification, together with an effective solver named GaSoliNe. Extensive experimental results demonstrate

Distribution Statement: 1-Approved for public release; distribution is unlimited.

Acknowledged Federal Support: Y

CONFERENCE PAPERS:

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: WWW '22: The ACM Web Conference 2022
Date Received: 04-Nov-2022 **Conference Date:** 01-May-2022 **Date Published:**
Conference Location: Virtual Event, Lyon France
Paper Title: Graph Sanitation with Application to Node Classification
Authors: Zhe Xu, Boxin Du, Hanghang Tong
Acknowledged Federal Support: Y

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: WWW '22: The ACM Web Conference 2022
Date Received: 04-Nov-2022 **Conference Date:** 01-May-2022 **Date Published:**
Conference Location: Virtual Event, Lyon France
Paper Title: Graph Communal Contrastive Learning
Authors: Bolian Li, Baoyu Jing, Hanghang Tong
Acknowledged Federal Support: Y

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: WWW '22: The ACM Web Conference 2022
Date Received: 04-Nov-2022 **Conference Date:** 01-May-2022 **Date Published:**
Conference Location: Virtual Event, Lyon France
Paper Title: Adversarial Graph Contrastive Learning with Information Regularization
Authors: Shengyu Feng, Baoyu Jing, Yada Zhu, Hanghang Tong
Acknowledged Federal Support: Y

RPPR Final Report as of 04-Nov-2022

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: WWW '22: The ACM Web Conference 2022
Date Received: 04-Nov-2022 Conference Date: 01-May-2022 Date Published:
Conference Location: Virtual Event, Lyon France
Paper Title: RawlsGCN: Towards Rawlsian Difference Principle on Graph Convolutional Network
Authors: Jian Kang, Yan Zhu, Yinglong Xia, Jiebo Luo, Hanghang Tong
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining
Date Received: 04-Nov-2022 Conference Date: 10-Aug-2022 Date Published:
Conference Location: Washington DC USA
Paper Title: Joint Knowledge Graph Completion and Question Answering
Authors: Lihui Liu, Boxin Du, Jiejun Xu, Yinglong Xia, Hanghang Tong
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining
Date Received: 04-Nov-2022 Conference Date: 10-Aug-2022 Date Published:
Conference Location: Washington DC USA
Paper Title: JuryGCN: Quantifying Jackknife Uncertainty on Graph Convolutional Networks
Authors: Jian Kang, Qinghai Zhou, Hanghang Tong
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: CIKM '22: The 31st ACM International Conference on Information and Knowledge Management
Date Received: 04-Nov-2022 Conference Date: 25-Oct-2022 Date Published:
Conference Location: Atlanta GA USA
Paper Title: Dissecting Cross-Layer Dependency Inference on Multi-Layered Inter-Dependent Networks
Authors: Yuchen Yan, Qinghai Zhou, Jinning Li, Tarek F. Abdelzaher, Hanghang Tong
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: CIKM '22: The 31st ACM International Conference on Information and Knowledge Management
Date Received: 04-Nov-2022 Conference Date: 25-Oct-2022 Date Published:
Conference Location: Atlanta GA USA
Paper Title: SuGeR: A Subgraph-based Graph Convolutional Network Method for Bundle Recommendation
Authors: Zhenning Zhang, Boxin Du, Hanghang Tong
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 3-Accepted
Conference Name: IEEE ICDM
Date Received: 04-Nov-2022 Conference Date: 01-Dec-2022 Date Published: 01-Nov-2022
Conference Location: Orlando, FL
Paper Title: Active Heterogeneous Graph Neural Networks with Per-step Meta-Q-Learning
Authors: Yuheng Zhang, Yinglong Xia, Yan Zhu, Yuejie Chi, Lei Ying, Hanghang Tong
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 04-Nov-2022

Publication Type: Conference Paper or Presentation **Publication Status:** 0-Other
Conference Name: IEEE ICDM
Date Received: 04-Nov-2022 Conference Date: 01-Dec-2022 Date Published: 01-Dec-2022
Conference Location: Orlando, FL
Paper Title: Generalized Few-Shot Node Classification
Authors: Zhe Xu, Kaize Ding, Yu-Xiong Wang, Huan Liu, Hanghang Tong
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 0-Other
Conference Name: NeulPS 2022
Date Received: Conference Date: 28-Nov-2022 Date Published: 28-Nov-2022
Conference Location: New Orleans, LA
Paper Title: CONSOLE: Convex Neural Symbolic Learning
Authors: Haoran Li, Yang Weng, Hanghang Tong
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 3-Accepted
Conference Name: NeulPS 2022
Date Received: 04-Nov-2022 Conference Date: 28-Nov-2022 Date Published: 28-Nov-2022
Conference Location: New Orleans, LA
Paper Title: Improved Algorithms for Neural Active Learning
Authors: Yikun Ban, Yuheng Zhang, Hanghang Tong, Arindam Banerjee, Jingrui He
Acknowledged Federal Support: **Y**

Partners

I certify that the information in the report is complete and accurate:
Signature: Hanghang Tong
Signature Date: 11/4/22 11:24AM

A GPU/CPU Hybrid Cluster for Accelerating Network-of-X Research and Education

Hanghang Tong, htong@illinois.edu, University of Illinois at Urbana-Champaign

In the age of AI and big data, networks provide powerful context to connect different things (e.g., entities, datasets, models) together. This phenomenon is referred to as Network-of-X in this project. By linking different entities, datasets, and data mining models together, Network-of-X brings a unique opportunity to dramatically enrich the modeling power, boost the mining performance, and broaden the reasoning capabilities.

The project aims to build a networked infrastructure to accelerate Network-of-X research and education. The proposed infrastructure has three mutually complementary modules. The first module (Distributed Graph Database) consists of a storage-optimized cluster offering high disk throughput and a distributed graph database management system supporting fast transaction processing. The second module (Graph Computation Engine) consists of a GPU server that is computationally optimized. It supports two types of fundamental computations in Network-of-X, including the large sparse matrix computations and the deep learning computations. The third module (Interactive Graph Analytics) develops an integrated human-computer interactive platform that allows the end-users to effectively collaborate and coordinate with computers in decision making, cognitive modeling, and algorithm training and diagnosis.

Outcome: Research

Using the built infrastructure from this project, including four dell PowerEdge C4140 GPU servers and 1 Dell 6.4TB NVMe Express Flash, we have advanced the research in network-of-X on several topics, including active graph neural networks, graph structure learning, graph contrastive learning, fair graph learning, knowledge graph reasoning, uncertain quantification on graphs, bundle recommendation, cross-layer dependence inference, equation learning, and few-shot learning on graphs. In total, it has led to 12 publications on these topics. These research outcomes are parts of the key deliverables in two DoD projects, including disinformation detection (DHS 17STQAC00001-03-03) and influence campaign and sense-making (DARPA INCAS HR001121C0165). In addition, the built infrastructure has benefited the following federal-sponsored projects, including network robustification (NSF 1947135), fair network learning (NSF 1939725), optimal deep graph learning (NSF 2134079), and AI for agriculture (NIFA 2020-67021-32799).

Outcome: Training and Education

In terms of education and training, we have used the built infrastructure as a development and testing platform for training 25 graduate students and 10 undergraduate students in various topics of large-scale graph mining and learning. During the project period, we graduated two Ph.D students (Dr. Si Zhang and Dr. Boxin Du) and one master's student (Zhe Xu). In addition, research opportunities were provided to both graduate and undergraduate students in the form of course projects in CS512: Data Mining Principles that PI taught in spring 2021. We also host a visiting student (Haoran Li) from Arizona State University. PI and his students have given a tutorial

on the topic of algorithmic fairness on network learning at ACM KDD 2022 and another tutorial on new frontier of multi-network mining at ACM KDD 2021. PI gave a plenary talk at Artificial Intelligence, Machine Learning and Data Science World Forum 2022 on the topic of “NetFair: Toward the Why Question of Network Mining”. He gave a keynote talk at Meta on the topic of “User Response Prediction: A Graph-based Approach” and another keynote talk at the Deep Learning on Graphs: Method and Applications (DLG-AAAI’21) on the topic of “NetFair: Towards Fair Network Mining”.

Dissemination

- In total, this project has led to 12 publications, which have been or will be presented at the top conferences in data mining and AI (e.g., KDD, WWW, AAAI, NeuIPS).
- PI and his students have given a tutorial on the topic of algorithmic fairness on network learning at ACM KDD 2022 and another tutorial on new frontier of multi-network mining at ACM KDD 2021.
- PI gave a plenary talk at Artificial Intelligence, Machine Learning and Data Science World Forum 2022 on the topic of “NetFair: Toward the Why Question of Network Mining”.
- PI gave a keynote talk at Meta on the topic of “User Response Prediction: A Graph-based Approach” and another keynote talk at the Deep Learning on Graphs: Method and Applications (DLG-AAAI’21) on the topic of “NetFair: Towards Fair Network Mining”.
- PI served as a panelist to discuss challenges and opportunities of the fairness in outlier detection at ODD workshop at KDD 2021.
- PI served as ACM CIKM2021 program committee co-chair.
- PI co-organized AI and social responsibility symposium.
- PI co-organized a workshop on Graph Techniques for Adversarial Activity Analytics at IEEE BigData 2021.

Awards/Honor

- PI received Dean's Award for Excellence in Research.
- PI was elevated to IEEE Fellow, for “contributions to graph mining”.
- PI received Excellent Paper Award from IEEE Big Data Mining and Analytics.
- PI received Best Student Paper Runner-up from IEEE ICDM 2021.
- PI received Visa Research Faculty Award.
- Master student Yuheng Zhang received “Kuck Outstanding Master’s Thesis Award”.
- Ph.D student Jian Kang was selected as a Mavis Future Faculty Fellow and was recognized as UChicago Rising Stars in Data Science.
- Master student Shengyu Feng was recognized as a Siebel scholar and received UIUC Graduate College Conference Presentation Award.
- Ph.D student Yuheng Zhang received UIUC Graduate College Conference Presentation Award.