| COVER SHEET FOR GRANT PROPOSALS<br>State Board of Education                           |                                    |   |            |  |  |
|---|------------------------------------|---|------------|--|--|
| SBOE PROPOSAL NUMBER:<br>(to be assigned by SBOE)                                     |                                    | AMOUNT REQUESTED: \$75,000              |            |  |  |
| TITLE OF PROPOSED PROJECT:<br>MAlicious Community Extractor (MAC<br>Criminal Networks | E): A Robust Toolkit for Unmasking |   |            |  |  |
| SPECIFIC PROJECT FOCUS:   |                                    |   |            |  |  |
| Cyberforensics, Cybersecurity   |                                    |   |            |  |  |
|   |                                    |   |            |  |  |
| PROJECT START DATE: July 1, 201   | 7                                  | PROJECT END DATE: June 30,              | 2018       |  |  |
| NAME OF INSTITUTION:<br>Boise State University  |                                    | DEPARTMENT: Computer Science Department |            |  |  |
| ADDRESS:1910 Unversity Dr., Boise   | , ID 83725-1135                    |   |            |  |  |
| E-MAIL ADDRESS: gabydagher@boisestate.edu   |                                    | PHONE NUMBER: (208) 426-57              | 782        |  |  |
|   | NAME:                              | TITLE:                                  | SIGNATURE: |  |  |
| PROJECT DIRECTOR/PRINCIPAL<br>INVESTIGATOR  | Gaby Dagher                        | Assistant Professor                     |            |  |  |
| CO-PRINCIPAL INVESTIGATOR   |                                    |   |            |  |  |
| NAME OF PARTNERING COMPANY:   |                                    | COMPANY REPRESENTATIVE NAME:            |            |  |  |
|   | NAME:                              | SIGNATURE:                              |            |  |  |
| Authorized Organizational<br>Representative   | Karen Henry                        | Horece                                  | Ala-       |  |  |
|   |                                    |   |            |  |  |
|   |                                    |   |            |  |  |

|  |                             |                             | SUMMARY PROP           | OSAL BUDGET |             |           |                         |  |
|--|-----------------------------|-----------------------------|------------------------|-------------|-------------|-----------|-------------------------|--|
| Name of Institution:<br>Name of Project Direct | tor:                        |                             |                        |             |             |           | 4                       |  |
| A. PERSONNEL COS<br>Graduate/Undergradua       | Г (Faculty, Staff,          | Visiting Professors<br>ner) | , Post-Doctoral Associ | ates,       |             |           |                         |  |
| Name/ Title                                    |                             |                             |                        | Salary/F    | Rate of Pay | Fringe    | Dollar Amount Requested |  |
| Gaby Dagher/ Assistar                          | nt Professor                |                             |                        | \$21,590    | )           | \$7,340   | \$28,930                |  |
| Graduate Student Assi                          | stant                       |                             |                        | \$24,000    | )           | \$4,429   | \$28,429                |  |
| Undergraduate Assista                          | nt                          |                             |                        | \$8,448     |             | \$439     | \$8,887                 |  |
|  |                             |                             |                        |             |             |           |                         |  |
|  |                             |                             |                        |             |             |           |                         |  |
|  |                             |                             |                        |             |             |           |                         |  |
|  |                             |                             |                        |             |             |           |                         |  |
| % OF TOT                                       | AL BUDGET:                  | 88%                         |                        |             | S           | SUBTOTAL: | \$66,246                |  |
| B. EQUIPMENT: (List<br>Item/D                  | each item with a escription | a cost in excess of         | \$1000.00.)            |             |             |           | Dollar Amount Requested |  |
|  |                             | 880.0 M                     |                        |             |             |           |                         |  |
|  |                             |                             |                        |             |             |           |                         |  |
|  |                             |                             |                        |             |             |           |                         |  |
|  |                             |                             |                        |             |             |           |                         |  |
|  |                             |                             |                        |             |             |           |                         |  |
| 5  |                             |                             |                        |             |             |           | ß                       |  |
|  |                             |                             |                        |             |             |           |                         |  |
|  |                             |                             |                        |             |             |           |                         |  |
|  |                             |                             |                        |             | SUBTO       | TAL:      |                         |  |
| C. TRAVEL:<br>Dates of Travel<br>(from/to)     | No. of<br>Persons           | Total<br>Days               | Transportation         | Lodging     | Per [       | Diem      | Dollar Amount Requested |  |
|  |                             |                             |                        |             |             |           |                         |  |
|  |                             |                             |                        |             |             |           |                         |  |
|  |                             | -                           |                        |             |             |           |                         |  |
|  |                             |                             |                        |             | SUBTOT      | TAL:      |                         |  |
| D. Participant Support                         | Costs:                      |                             |                        |             |             |           | Dollar Amount Requested |  |
| 1. Stipends                                    |                             |                             |                        |             |             | \$8       | ,754                    |  |
| 2. Other                                       |                             |                             |                        |             |             |           |                         |  |
|  |                             |                             |                        |             | SUBTOTA     | AL:       | ,754                    |  |

| E. Other Direct Costs:                               |           | Dollar Amount Requested |
|--|-----------|-------------------------|
| 1. Materials and Supplies                            |           |                         |
| 2. Publication Costs/Page Charges                    |           |                         |
| 3. Consultant Services (Include Travel Expenses)     |           |                         |
| 4. Computer Services                                 |           |                         |
| 5. Subcontracts                                      |           |                         |
| 6. Other (specify nature & breakdown if over \$1000) |           |                         |
|  |           |                         |
|  | SUBTOTAL: |                         |
|  |           | \$75,000                |
| F Total Costs: (Add subtotals, sections A through E) | TOTAL:    |                         |
|  |           |                         |
|  |           | \$75,000                |
| G Amount Requested:                                  | TOTAL:    |                         |
| Project Director's Signature:                        | Date:     |                         |
|  |           |                         |

| INSTITUTIONAL AND OTHER SE<br>(add additional pages as r |        |
|--|--------|
| A. INSTITUTIONAL / OTHER SECTOR DOLLARS                  |        |
| Source / Description                                     | Amount |
|  |        |
|  |        |
| B. FACULTY / STAFF POSITIONS                             |        |
| Description  |        |
|  |        |
|  |        |
| C. CAPITAL EQUIPMENT                                     |        |
| Description  |        |
|  |        |
|  |        |
| D. FACILITIES & INSTRUMENTATION (Description)            |        |
|  |        |
|  |        |
|  |        |

# MAlicious Community Extractor (MACE): A Robust Toolkit for Unmasking Criminal Networks

| 1 | INSTITUTION              | Boise State University   |
|---|--------------------------|--|
| 2 | FACULTY MEMBER DIRECTING | Dr. Gaby Dagher, Assistant Professor, Computer Science,<br>Information Security, Privacy, and Mining (ISPM) Research Lab |

### 3. Award Status

A portion of the technology leveraged for this proposal was awarded incubation funding last year. The project, Cyber Forensic Investigation Toolkit (CFIT): Next-generation Evidence Gathering for Law Enforcement, was awarded \$58,000. This proposal builds on the CFIT technology developed from previous work to create a cyber forensics software suite that correlates diverse information from large data sets and graphically visualizes relationships between people and topics. This new cyber forensics software suite, *MAlicious Community Extractor (MACE)*, is uniquely positioned to provide law enforcement agencies a much needed tool to identify criminal networks.

## 4. Executive Summary

During criminal investigations, being able to efficiently identify, collect and analyze digital evidence from a suspect's computer is critical to the development and prosecution of criminal cases. In cases involving organized crime, terrorism, conspiracy, aiding and abetting or multiple principals, determining the extent of criminal networks and their communication patterns is very valuable to law enforcement agencies. However, correlating information from an exceedingly large number of files that exist through emails, texts, text documents and social media is only possible using manual, time-intensive techniques that are prone to human error. To automate criminal network detection from large data sets, we propose the refinement and commercialization of the *MAlicious Community Extractor (MACE)* toolkit. MACE is designed using a novel

sequential pattern mining algorithm that efficiently identifies potential criminal networks while providing an interactive interface that graphically displays communication patterns and relationships within and between the criminal networks. As proof of concept, the *Cyber Forensic Unit (CFU)* at the *Information Security, Privacy, and Mining (ISPM) Research Lab,* which is directed by Dr. Dagher, has implemented all these core features and recently built an experimental alpha prototype to showcase the enhanced functionality available with this new investigation toolkit.

## 5. Total Amount Requested to Meet "Gap" Project Objectives

To meet the following project objectives, our team requests \$75,000. We will:

- Design the existing MACE prototype for commercialization. To design MACE into a reliable, fully-fledged commercial product for law enforcement agencies, we will focus on:
  - <u>Investigator-Centered Interface</u>: We have developed a functional graphical user interface for MACE. However, the interface needs refining to align with the needs of investigators. Given the letter of support from the Idaho State Police (ISP), our MACE team will work closely with investigators to customize the visualization framework and implement an investigator-centered interface that is optimized for how investigators intend to use the MACE toolkit.
  - <u>Quality Assurance</u>: We will design test cases, define clear quality measures, and construct and execute a quality assurance plan to thoroughly test the MACE toolkit, including the new user interface.

- <u>Benchmarking</u>: We will benchmark MACE performance against existing state-of-the art cyber forensic tools, including Forensic Toolkit® by AccessData Group, Inc., and EnCase® by Guidance Software, Inc.
- <u>Beta Phase Testing</u>: We will work with the ISP to test the MACE toolkit using ISP-sourced data to evaluate usability, to ensure functionality and to validate accuracy.
- 2) Market MACE to cybersecurity companies and to law enforcement agencies at the local, state, and federal levels. We plan to prepare a rigorous marketing plan that includes constructing a list of prospective customers, clearly defining marketing goals, and determining marketing communication strategies and tactics.

## 6. Resource Alignment with Boise State University Priorities

As part of the 2012-2017 strategic plan toward becoming a *metropolitan research university of distinction*, Boise State University has placed great emphasis on STEM disciplines. In recent years, the number of students majoring in STEM disciplines increased 66% while overall growth in the student body was 5%. The emphasis placed on STEM disciplines and research productivity led to Boise State being designated as a doctoral research institution by the Carnegie Classification of Institutions of Higher Education in 2016. This achievement highlights Boise State's commitment to developing research intensive programs, which is underscored by the establishment of the Computing Ph.D. Program last fall. And with a cybersecurity emphasis as part of the Computing Ph.D. Program, this project is in direct alignment with Boise State's priorities.

## 7. Specific Project Plan and Budget

## 7a. The Market Opportunity

The MACE toolkit has enormous potential, as demonstrated by the market need, demand, and audience.

Market Needs. A central element of MACE is to identify criminal networks and define communities (groups of people that a suspect frequently interacts with). However, this task is daunting due to the large amount of information contained within digital devices. The complexity of identifying criminal networks within unstructured text documents – such as text documents, text from social media platforms, and emails – becomes even more involved with the continuously increasing capacity of data storage devices. There are two major problems with today's digital forensic investigation (DFI) tools:

**1. Limited Capability.** Existing DFI tools allow for various types of searches, including keyword, regular expression, and approximation matching; which in turn allow the user to extract distinct information from a suspect's computer, such as names, phone numbers, and addresses. These tools are incapable of performing complex data mining techniques needed to identify distinct topics and relate them to criminal networks.

**2.** Pattern of Communication. In recent years, several approaches [1][2] were proposed in the literature to detect criminal networks from a given document set contained in digital storage devices. However, none of these approaches can or even attempt to determine the pattern of communication among members of a network, which is essential in investigations where unnamed accomplices may be involved.

Applications and Markets for the Technology. Out of all types of available data in cyber forensic investigations, text-based data is the most common medium for communication between individuals. However, this type of data is also the most challenging to analyze, as it is not a trivial task to program a software system to automatically interpret the contextual meaning of phrases.

The MACE toolkit performs semantic analysis of phrases in all documents. Based from the semantic analysis, MACE determines distinct topics embedded in the documents, identifies criminal networks per topic, and allows investigators to efficiently visualize (1) the networks in a very large set of documents, (2) the relationships between the networks, and (3) the pattern of communication within each network.

Potential market audience, competition, and barriers to market entry. The market for a MACE toolkit is broad, as any law enforcement agency interested in extracting criminal networks from a suspect's digital media could employ it, and for all types of crimes. In addition, cybersecurity companies can also utilize the proposed technology to extract relevant information about on-line malicious attacks and breaches carried out against a terminal machine or server. Several cyber forensic tools are commercially available, including Forensic Toolkit®, EnCase®, CAINE and The Sleuth Kit®. However, unlike the MACE toolkit, there is currently no theoretical approach nor commercial tool that allows the investigator to visualize the criminal networks and their pattern of communications in documents captured on a suspect's computer. As a result, we expect law enforcement agencies as well as cybersecurity companies will be eager to add MACE to their cyber forensics laboratories.

## 7b. The Technology and Path to Commercialization

The team at the *ISPM Research Lab* has built a working prototype that includes all core functionalities: a graphical user interface, forensics analysis, data mining, and a visual output. The technology, while mature, has never been integrated in such a way that would enable a simple, yet effective data mining/cyber forensics analysis tool. The forensics and data mining backbone of MACE is based on Java and R Data Mining programs, while the visualization framework was built

using D3 library, an open source JavaScript library for visualization. To transform the MACE prototype to a commercial grade product, it needs to be further developed to include a user-friendly, investigator-centered interface. Once benchmarked to ensure it meets industry quality and performance standards, a MACE toolkit prototype will undergo beta phase testing.

The MACE technology. The proposed MACE toolkit consists of three main components:

**1) Algorithm for Extracting Communities (Algorithm 1).** The document set is first clustered based on topics using a bisecting k-means algorithm, and then a frequent pattern mining algorithm is applied to determine the individuals who frequently correspond with each other on the same topic. This algorithm also determines the relations between communities.

**2**) Algorithm for Determining Pattern of Communications (Algorithm 2). Utilizing advanced techniques in sequential pattern mining, we have developed a robust algorithm to determine in each criminal network the pattern of communication (PoC) among its members, as illustrated in Figure 1. MACE is the first cyber forensic tool that investigates and extracts patterns of communication from malicious communities.



Figure 1: Communication patterns within a malicious community.

**3) Visualization Framework.** Based on the output from Algorithm 1 and 2, an interactive interface allows an investigator to visualize the communities, their relations, and their patterns of communications, as shown in Figure 2. The gray area shows the six communities extracted, and

the right sidebar lists details about the selected community (i.e., the green community with outlined in red).



Figure 2: MACE visualization framework showing different communities extracted from a given document set. Market need for MACE technology and its intellectual property status. As indicated by the letter of support from the ISP, there is a need for tools such as the MACE toolkit. The intellectual property of the MACE toolkit belongs to Boise State University since the PI used university funds and facilities to develop it. We will work with the Office of Technology Transfer at BSU to patent the technology--System and method for extracting malicious networks and establishing the pattern of communications within each network."

Who developed the technology and with what funding? Dr. Gaby Dagher, faculty in the Department of Computer Science at Boise State University worked with his students to build the MACE prototype. A university graduate assistantship provided funding.

Concrete steps to bring MACE technology to market.

In parallel to refining the MACE prototype for commercialization, we plan to execute a rigorous market plan which will inform our decision on how to best capitalize on the technology described within this proposal. We will market MACE to cybersecurity companies, and to law enforcement agencies at the local, state, and federal levels.

Dependent upon the results from our rigorous market analysis and planning, we intend to effectuate the commercialization of the MACE toolkit by creating a start-up company or directly outlicensing the technology to an established company for dissemination.

- The PI will work with the Office of Technology Transfer (OTT) in a two-pronged approach for the most effective and appropriate licensing strategy: (1) creating a startup around this technology or (2) out-licensing to an existing entity.
  - 2. The PI will complete the internal process for Entrepreneurial Conflicts of Interest for startup formation while concurrently identifying potential licensees.
  - 3. In preparing the licensing strategy, the PI and the OTT will analyze each licensing opportunity individually in a manner that reflects the business needs and values of BSU.

Pursuant to the OSBE Licensing Guidelines, many aspects must be considered such as: "the nature and stage of development of the technology; the breadth and complexity of the potential fields of use; the product development path and timeline; the extent of intellectual property protection; the relevant markets and market niches; specific campus practices; unique needs of prospective licensees; ethical considerations for the use of future products; and emerging issues, among other elements."<sup>1</sup> Most importantly, we would seek "licensees capable of bringing the invention to the marketplace in a timely manner. While often only one potential licensee comes forward for any

<sup>&</sup>lt;sup>1</sup> Idaho State Board of Education, Institution Technology Licensing Guidelines, June 2013

given institution invention, the institution should nevertheless assess the potential licensee's technical, managerial and financial capability to commercialize the technology."<sup>1</sup> As such, while our preference would be to license this technology to a faculty startup, we must nonetheless consider all licensing plays for timely dissemination of the technology and the business needs of the University.

## 7c. Commercialization Partners

Given that currently there are no cyber forensics company based in Boise, it is a great opportunity for us to lay the foundational work for the industry. By commercializing through a start-up or licensing, we will have the opportunity to develop relationships with multiple partners within the state, including law enforcement agencies, INL, and businesses. Given the supporting letter from ISP, the department will collaborate with the *ISPM Research Lab* team throughout the project and support our effort to build a commercial version of the tool. The below outlines collaboration detail.

| ENTITIES | RESPONSIBLE PROJECT TASKS  |
|----------|--|
| PI + ISP | Design Investigator-Centered Interface                                   |
| PIs      | Develop Investigator-Centered Interface, Quality Assurance, Benchmarking |
| PI + ISP | Beta Phase Testing   |

# 7d. Institutional and Other Support

The computational tools required for this research are provided by Boise State University. The collaboration with the Idaho State Police will provide functional and test feedback for product development and benchmarking.

# 7e. Budget, Project Plan and Fund Use

We will not use grant funds to purchase equipment or pay for travel.

| FOR                   | DESCRIPTION  | BUDGET   |
|-----------------------|--|----------|
| Dr. Gaby Dagher (PI)  | Two months of summer salary and fringe benefit for research and oversight      | \$28,930 |
| Graduate Student      | Salary, fringe benefits, health insurance, and tuition for 1 year              | \$37,183 |
| Undergraduate Student | Salary and fringe benefits at \$13 per hour and 10 hours per week for 52 weeks | \$8,887  |
|                       | Total  | \$75,000 |

Here is the project timeline with milestones highlighted: (A): User Interface, (B): Quality

Assurance and Benchmarking, and (C): Experimental Evaluation.

| # | PROJECT TASKS                    | Q1 | Q2 | Q3 | Q4 |
|---|----------------------------------|----|----|----|----|
|   | Milestones:                      |    | А  | В  | С  |
| 1 | User-Centered Interface          | Х  | Х  |    |    |
| 2 | Quality Assurance + Benchmarking |    | х  | Х  |    |
| 3 | Beta Phase Testing               |    |    | Х  | Х  |

PI Dagher will oversee the project budget, coordinate with ISP, and be responsible for ensuring project reports and deliverables are completed on time. The graduate student will handle the design and development of the investigator-centered interface, as well as beta phase testing. The undergraduate student will be responsible for quality assurance and benchmarking.

## **References:**

[1] Al-Zaidy, Rabeah, Benjamin C. M. Fung, Amr M. Youssef and Francis Fortin, 'Mining criminal networks from unstructured text documents', Digital Investigation, 2012.

[2] Lau, R.Y.K., Xia, Y. and Ye, Y. (2014) A Probabilistic Generative Model for MiningCybercriminal Networks from Online Social Media. IEEE Computational Intelligence Magazine.

## Facilities and Equipment

The Computer Science Department currently occupies a new facility located in the center of downtown Boise. Located approximately 0.5 miles from the Boise State University campus, the new facility is within easy walking distance, and is served by a free shuttle bus system that is available on an 8 minute frequency. This new downtown building also includes the main transportation hub for the entire Boise City bus system, truly locating Computer Science at the center of Boise.

The Computer Science department will occupy the second and third floor of the City Center Plaza, is a 9 story building, with a total footprint of 53,549 GSF. The downtown area is home to a large number of software development firms, including one of the largest in Boise who will be co-located in the upper floors of the building, providing a unique opportunity for the collaboration between industry and students. The new facility includes server rooms, a visualization center, tutoring center, 34 faculty/staff and departmental offices, 6 classrooms, conference rooms, focus rooms, graduate student offices, outdoor balconies and community meeting and gathering spaces. The facility is connected to the university core on dark fiber. The initial connection is 10 gig, scalable to meet future needs with additional optics. One gig connection is available to all connected network devices at the site. Wireless connectivity is provided through dense 2.4 ghz and 5 ghz wireless AC radios.

### Computer Science Servers

The CS department-maintained MEC 305 is a climate controlled server room which houses the Beowulf Cluster Lab. This lab was originally funded by a NSF MRI grant. Since then, it has

received additional funding from the DoD, FAA, NASA, and several private companies. Currently, the lab has four clusters with approximately 184 processors, 500 GB of RAM, and over 100 TB of storage. The GeneSIS storage cluster, a 21 node Beowulf style cluster. Two of the clusters run Hadoop for the study of big data. One cluster utilizes 56Gb/s Infiniband. The other cluster uses traditional Ethernet connections. Both clusters run Apache Ambari for management. The dedicated Beowulf style processing cluster has 58 nodes supporting Intel dual core Xeon processors.

#### **Computer Science Teaching and Research Labs**

The Computer Science department maintains a 32 workstation Linux lab for students. This lab has special software and hardware to support advanced Computer Science courses, including Microsoft Windows which is available through virtualization. This laboratory is sponsored by MetaGeek, a Boise-based software company that is a leader in the field of wireless network analysis software. The Computer Science Department also maintains a tutoring center that houses 30 computers and 1 teaching station. This lab has Linux software that supports the introductory CS courses. The Department employs approximately 20 undergraduate lab assistants and graduate teaching assistants each semester who support of a variety of courses. The tutoring center was established with support from the IGEM grant.

## Biographical Sketch

# GABY DAGHER, PH.D.

#### PROFESSIONAL PREPARATION

| INSTITUTION          | LOCATION         | MAJOR                                      | DEGREE & YEAR |
|----------------------|------------------|--|---------------|
| Damascus University  | Damascus, Syria  | Engineering Management<br>and Construction | B.S., 1993    |
| Concordia University | Montreal, Canada | Computer Science                           | B.S., 2002    |
| Concordia University | Montreal, Canada | Information Systems<br>Security            | M.S., 2011    |
| Concordia University | Montreal, Canada | Computer Science                           | Ph.D., 2015   |

#### APPOINTMENTS

| PERIOD       | APPOINTMENT         | INSTITUTION & LOCATION   |
|--------------|---------------------|--|
| 2016–Present | Assistant Professor | Computer Science, Boise State University, Boise ID                   |
| 2014–2016    | Research Assistant  | Data Mining and Security Lab, McGill University,<br>Montreal, Canada |
| 2013-2015    | Instructor          | Computer Science, Concordia University, Montreal, Canada             |
| 2011-2015    | Research Assistant  | CIISE, Concordia University, Montreal, Canada                        |

#### PRODUCTS

PROJECT-RELATED

 Gaby G. Dagher, Benedikt Bünz, Joseph Bonneau, Jeremy Clark, and Dan Boneh. Provisions: Private Proofs of Solvency for Bitcoin Exchanges. In 22<sup>nd</sup> ACM Conference on Computer and Communications Security (CCS), 12 pages, 2015.
[Link: http://cs.boisestate.edu/~gdagher/pub/GD Provisions CCS 2015.pdf]

[Link: <u>http://cs.boisestate.edu/~gdagher/pub/GD\_Provisions\_CCS\_2015.pdf</u>]

- Gaby G. Dagher, Farkhund Iqbal, Mahtab Arafati and Benjamin. C. M. Fung. Fusion: Privacypreserving Distributed Protocol for High-Dimensional Data Mashup. In 21<sup>st</sup> IEEE International Conference on Parallel and Distributed Systems (ICPADS), 10 pages, 2015.
  [Link: <u>http://cs.boisestate.edu/~gdagher/pub/GD\_Fusion\_ICPADS\_2015.pdf</u>]
- <u>Mahtab Arafati</u>, Gaby G. Dagher, Benjamin C. M. Fung, and Patrick C. K. Hung. D-Mash: A Framework for Privacy-Preserving Data-as-a-Service Mashups. In CLOUD, 8 pages, 2014.
  [Link: <u>http://cs.boisestate.edu/~gdagher/pub/GD\_DMash\_CLOUD\_2014.pdf</u>]
- Omar Abdel Wahab, Moulay Omar Hachami, Arslan Zaffari, Mery Vivas, and Gaby G. Dagher. DARM: A Privacy-preserving Approach for Distributed Association Rules Mining on Horizontally-partitioned Data. In IDEAS, 8 pages, 2014.
  [Link: http://cs.boisestate.edu/~gdagher/pub/GD\_DARM\_IDEAS\_2014.pdf]
- Gaby G. Dagher and Benjamin. C. M. Fung. Subject-based Semantic Document Clustering for Digital Forensic Investigations. Data & Knowledge Engineering (DKE), Elsevier, vol. 86, 2013.
  [Link: <u>http://cs.boisestate.edu/~gdagher/pub/GD\_ClusterDFI\_DKE\_2013.pdf</u>]

- Junnan Chen, Courtney Miller, and Gaby G. Dagher. Product Recommendation System for Small Online Retailers Using Association Rules Mining. In ICIDM, 6 pages, 2014.
  [Link: <u>http://cs.boisestate.edu/~gdagher/pub/GD\_PRS\_ICIDM\_2014.pdf</u>]
- Gaby G. Dagher, Benjamin. C. M. Fung, Noman Mohammed, and Jeremy Clark. SecDM: A Privacy-preserving Framework for Confidential Query Processing on the Cloud. Knowledge and Information Systems (KAIS). 20 pages. Under second revision.
  [Link: <u>http://cs.boisestate.edu/~gdagher/pub/GD\_SecDM\_KAIS\_2016.pdf</u>]

### SYNERGISTIC ACTIVITIES

### **Review Articles:**

Reviewing manuscripts for International Conference on Knowledge Discovery and Data Mining (SIGKDD), IEEE International Conference on Big Data (BigData), Information Systems Frontiers Journal (ISFJ), IEEE International Conference on Data Engineering (ICDE), Proceedings of Very Large DataBases (PVLDB), ACM Transactions on Database Systems (TODS), IEEE International Conference on Services Computing (SCC), International Conference on Information Science, Signal Processing and their Applications (ISSPA), IEEE International Conference on Data Mining (ICDM), Elsevier Data & Knowledge Engineering (DKE).

### **Student Advising:**

Ishita Dwivedi, Master's student in Computer Science, Boise State University. Joshua Holmes, Master's student in Computer Science, Boise State University. Manish Kumar, Master's student in Computer Science, Boise State University. Anthony Harris, Master's student in Math, Boise State University.

### Grants Awarded:

- Idaho SBOE HERC Incubation Fund, "Cyber Forensic Investigation Toolkit (CFIT): Next-generation Evidence-gathering for Law Enforcement," PI: Gaby Dagher, Co-PI: Jyh-haw Yeh. 7/01/2016-6/30/2017, \$58,000.
- Wider Persist, Boise State University, "Applying Team-Based Learning in Applied Cryptography Graduate Course," PI: Gaby Dagher. 6/01/2016 12/15/2016, \$7,500.



Idaho State Police

Service Since 1939



C.L. "Butch" Otter Governor

Colonel Ralph W. Powell Director

March 27, 2017

State Higher Education Research Council Idaho State Board of Education 650 West State Street Boise, ID 83702

To Whom It May Concern,

The Idaho State Police is very supportive of the efforts of Idaho's institutions of higher education in developing tools to assist law enforcement in protecting the public and solving crimes. The efforts of Boise State University and Dr. Gaby Dagher in developing a toolkit to assist law enforcement with cyber forensic investigations is exciting and could be a true benefit to law enforcement, not only in Idaho but nationwide.

The proliferation of computers, smart phones, tablets and other portable digital devices are intertwined in our daily lives and are largely the way people communicate. This is no different for those operating in the criminal element. Criminals use computers, mobile phones and devices to further their criminal acts. Digital evidence is critical in most criminal cases from drug trafficking and financial crimes to violent crimes such as homicides, rapes, and child sexual abuse.

The gathering and analysis of digital information is vital to the prosecution of those committing these crimes. Law enforcement is allowed to gather this type of evidence with the permission of the court. Our agency alone serves approximately 700 search warrants each year on electronic devices.

A toolkit for use by law enforcement to analyze digital information collected from a suspect's computer, mobile phone, disc or flash drive would be valuable. The information collected can be used to detect hidden criminal networks and understand their order of communication, which could lead to the capture of other criminals, as well as the discovery of plans of future crimes that police can now prevent.

The Idaho State Police support the efforts of Boise State University and Dr. Gaby Dagher in securing a grant from the State Higher Education Research Council to develop a toolkit to assist in cyber forensic investigation.

Sincerely, Turell

Colonel Ralph W. Powell Director

700 South Stratford Drive • Meridian, Idaho 83642-6202