FACTA UNIVERSITATIS
Series: **Electronics and Energetics** Vol. 29, N° 3, September 2016, pp. 437 - 450
DOI: 10.2298/FUEE1603437L

A HADOOP-ENABLED SENSOR-ORIENTED INFORMATION SYSTEM FOR KNOWLEDGE DISCOVERY ABOUT TARGET-OF-INTEREST

Yu Liang, Chao Wu

Department of Computer Science and Engineering,
University of Tennessee at Chattanooga, USA

Abstract. *To obtain a real-time situational awareness about the specific behavior of targets-of-interest using large-scale sensory data-set, this paper presents a generic sensor-oriented information system based on Hadoop Ecosystem, which is denoted as SOIS-Hadoop for simplicity. Robotic heterogeneous sensor nodes bound by wireless sensor network are used to track things-of-interest. Hadoop Ecosystem enables highly scalable and fault-tolerant acquisition, fusion and storage, retrieval, and processing of sensory data. In addition, SOIS-Hadoop employs temporally and spatially dependent mathematical model to formulate the expected behavior of targets-of-interest, based on which the observed behavior of targets can be analyzed and evaluated. Using two real-world sensor-oriented information processing and analysis problems as examples, the mechanism of SOIS-Hadoop is also presented and validated in detail.*

Key words: *Sensor-oriented information system, Hadoop Ecosystem, target of interest, wireless sensor network, mathematical model.*

1. INTRODUCTION

An information system is generally a computer-centric system that integrates data acquisition, processing and analysis, storage and communication, interpretation, and knowledge discovery, etc. [1-4]. Sensor-oriented information systems addressed in this work aim to obtain a panoramic, timely, trusted understanding about the observed behavior of targets-of-interest (TOI) [3][5-10] by exploiting networked sensor assets, which consist of large number of autonomous, heterogeneous, and multi-layer sensor nodes. It is an extremely computationally intense, labor-intensive and highly unreliable job to derive a real-time situational awareness about TOI from high volume, high generation velocity, wide variety of sensory data, the addressed sensor-oriented information system is constructed over Hadoop Ecosystem [8][11], a conventionally used big-data platform.

Received July 1, 2015; received in revised form November 16, 2015
Corresponding author: Yu Liang
Department of Computer Science and Engineering, University of Tennessee at Chattanooga, Tennessee, 37403, USA
(e-mail: Yu-Liang@utc.edu)

This paper proposes a generic sensor-oriented information system based on Hadoop cluster (SOIS-Hadoop) [11-14] to monitor and analyze the specific behavior of target-of-interest (TOI) according to persistent surveillance sensory data. The addressed SOIS-Hadoop has the following features: (1) Employing temporally and spatially dependent mathematical models [3][9][15][16] to formulate the expected behavior about targets-of-interest, based on which the observed behavior of TOI can be evaluated or the future behavior of TOI can be predicted; (2) Tracking TOI through deploying networked autonomous sensor nodes, which will be tuned using collective control and self-optimization to achieve the optimal, reliable, and energy-efficient observations; (3) Using Hadoop Ecosystem to handle the acquisition, fusion, storage, management and mining of large-scale real-time/historical sensory data.

The major topics of this paper include: (1) a generic hardware and software infrastructure, and the implementation flowchart of SOIS-Hadoop; (2) the application of SOIS-Hadoop in real-world sensor-enabled engineering problems such as predictive analysis of the aggregation of carp, and the anomaly detection of traffic flow; (3) mathematical modeling about the expected behavior of TOI, both microscopic and macroscopic strategies are addressed.

The remainder of the paper is organized as follows: Section 2 discusses the hardware infrastructure of the SOIS-Hadoop system; Section 3 discusses the software framework of the SOIS-Hadoop system; Section 4 briefly introduces the flowchart of the system; Section 5 uses two representative sensor-oriented information analysis problems as benchmark to demonstrate the mechanism and implementation of SOIS-Hadoop; Section 6 introduces macro-cell strategy, a divide-and-conquer method, to manipulate large-scale application problems with high scalability; Section 7 summarizes the effort.

2. HARDWARE ARCHITECTURE OF HADOOP XEN CLUSTERS

The overarching goal of this interdisciplinary project is enabling robust intelligent systems which can operate autonomously for long periods of time. This ability requires that all system components are seamlessly integrated. Figure 1 shows the hardware-configuration of Hadoop based sensory data processing and analysis system. The proposed system consists of three hardware modules:

(1) Data acquisition and pre-processing based on mobile computing platform (e.g., iPhone, laptop, etc.). Sensory data may be acquired by multiple sensors at the same time. Pre-processing indicates translating stream sensory data such as video data into semi-structured format data such as XML format (www.w3.org).
(2) Data storage and management server based on Hadoop cluster, which is built using multiple inter-connected Xen virtual machines.
(3) Data analysis and visualization client, which mainly simulates the temporally and spatially dependent mathematical model.

To fully exploit sensor asset, the following issues need to be investigated: (1) the design and development of provably correct, decentralized algorithms for finding and localizing multiple mobile TOI using a networked sensor nodes [17], and synchronizing sensor stream, etc.; (2) achieving longevity and energy-efficiency by developing energy efficient motion planning algorithms; studying system level energy trade-offs and optimization; improving system life-time by energy harvesting; (3) mobility and energy aware communication protocols for robotic networks; and (4) data analysis algorithms to discover TOI's mobility patterns at multiple scales and how these patterns correlate with environmental parameters.

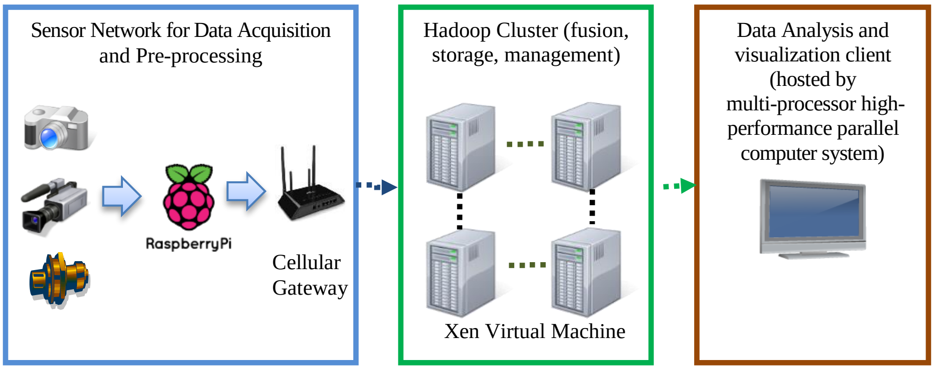

Fig. 1 Hardware Configuration of SOIS-Hadoop (dash-line indicates those internet link).

As illustrated in Figure 1, a multitude of collaborative mobile sensor nodes are deployed to detect, discriminate, localize, and track targets of interest (TOIs). Each sensor node is managed by a Raspberry Pi single-board computer (www.raspberrypi.org), which is equipped with the Robot Operating System (www.ros.org), 3G/4G CDMA cellular gateway that will provides sensor nodes with internet connectivity based on radio transceiver, solar panel, GPS-aided inertial navigation system [18], and electro-optical sensor used to capture the movement of TOIs. Robot Operating System provides standard operating system services such as low-level device control, implementation of commonly-used functionality, inter-process message passing, and package management.

Fig. 2 (a) Monitoring Invasive fish with an robotic boat (Univ. of Minnesota); (b)-(c) Tracking and analysis of the movement of vehicle according to low-resolution video data acquired by UAV flying over the city

Figures 2(A)-(C) illustrates that mobile sensor nodes [19] are deployed to track carp in lakes and monitor traffic status over a city respectively. Long-term operation of sensor nodes necessitates a long-term energy source. In this work, the energy consumption [20] for wireless sensor network [21] will be minimized from the point of view of mobility, communication, and solar harvesting. First, accurate trajectory estimation about moving

TOI will greatly optimize the motion of sensor-nodes; second, wireless communication is the largest energy consumer on the robotic platform, IEEE 802.15.4 wireless personal area networks protocol (standards.ieee.org) is used in our work; the topology of wireless sensor network, synchronization technique, WSN routing algorithm, transmission protocol, and long path-loss models [22] may all determine the energy efficiency of WSN. Third, the robotic vehicle for TOI tracking is equipped with solar panel coupled to a solar charger and a deep-cycle rechargeable battery.

A Hadoop ecosystem built on Xen Linux virtualization (www.xenproject.org) cluster provides a highly scalable and fault-tolerant platform for acquiring, fusing, storing and analyzing huge amounts of sensory data in a distributed computing environment.

Data analysis and visualization client mainly handle the simulation of temporally and spatially dependent mathematical model, the most computationally intensive operation in the implementation SOIS-Hadoop. In this work, Data analysis and visualization client is hosted by a multi-processor and multi-core parallel computing machine. Message Passing Interface (MPI) parallel programming paradigm is used to implement the simulation of temporally and spatially dependent mathematical model [8][23].

3. SOFTWARE INFRASTRUCTURE OF SOIS-HADOOP

Figure 3 demonstrates the software infrastructure about SOIS-Hadoop. Considering the addressed information system is driven by huge-scale heterogeneous sensory data [3], highly scalable, robust, and relatively accurate computational methods are investigated in the implementation of SOIS-Hadoop.

Accessory information (e.g., geographic information, weather condition, and historical data, which are of XML format in our implementation) system and persistent surveillance sensory data constitute two major important inputs for SOIS-Hadoop. Hadoop-Ecosystem [14] is employed as the main engine of SOIS-Hadoop: Flume (flume.apache.org) acquires, aggregates, pre-processes, and then forward the sensory data to Hadoop Distributed File System (HDFS); Sqoop (sqoop.apache.org) provides an interface between accessory information and HDFS; Built on the top of HDFS, HBase (hbase.aparche.org) provides a real-time and random access to the data; HBase is equipped with NoSQL Database [7], which is of key-value format, supports highly scalable, concurrent, and fault-tolerant storage about structured or semi-structured data appeared in SOIS-Hadoop because it does not need to category and parse the sensory data into fixed format; Hive (hive.apache.org) facilitates query and managing large dataset; R-connector and Mahout (mahout.apache.com) are employed in the mining and statistical analysis about sensory and accessory data. As one of our major contributions, analytics module extracts the features about target-of-interest from data using temporally and spatially dependent mathematical model. Classification and clustering module use machine learning strategy to measure the event according to the features obtained in Analytics module.

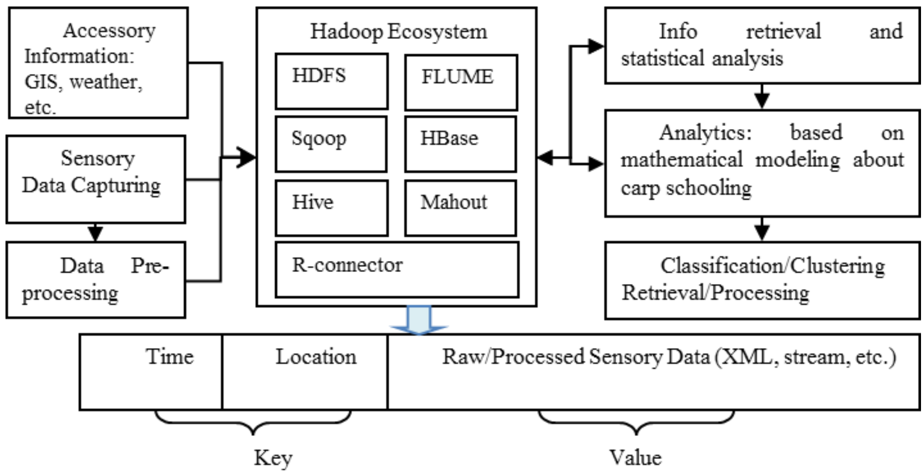

Fig. 3 Infrastructure for SOIS-Hadoop.

4. A Generic Flowchart for Sensor-oriented Information Analysis System

Fig. 4 shows a generic flow-chart about the sensor-oriented information analysis system [12]. Geographic information module defines the geometry configuration of the scene; mathematical model about the expected behavior of target-of-interest [6], which is emphatically investigated in this paper, agent-based mathematical model is used to anticipate the evolution of observed behavior about carp school; persistent surveillance sensory data is directly acquired from sensors. In addition, the acquisition, pre-processing, storage, retrieval are all implemented based on Hadoop ecosystem [11] including Apache Flume, Mahout, Hive, and R-connector, etc.

Fig. 4 also illustrates that the implementation of sensor-oriented information analysis system consists of following two threads: (1) Formulating the mathematical model (e.g., spatial- and temporal-dependent partial differential equations) using historical sensory data about the expected behavior of thing-of-interest (TOI) [5][6][13][24]. (2) Processing and integration of observed sensory data. A situational awareness is derived from these two threads. Then the situational awareness will reversely guide the self-optimization of data-collection and cooperative control [25] of sensor nodes so as to generate more accurate understanding about external events and achieve longevity by developing energy-efficient motion planning algorithms.

The mathematical model about the expected behavior of target-of-interest is formulated according to the geographic information and the historical records about target-of-interests. The core calculations corresponding to mathematics modeling include (1) statistical analysis about the historical sensory data stored in Hadoop Distributed File System (HDFS), and (2) numerical solution to the mathematical simulation (e.g., using finite element method [9][14][16][23] to solve the temporal and spatial-dependent partial differential equations.

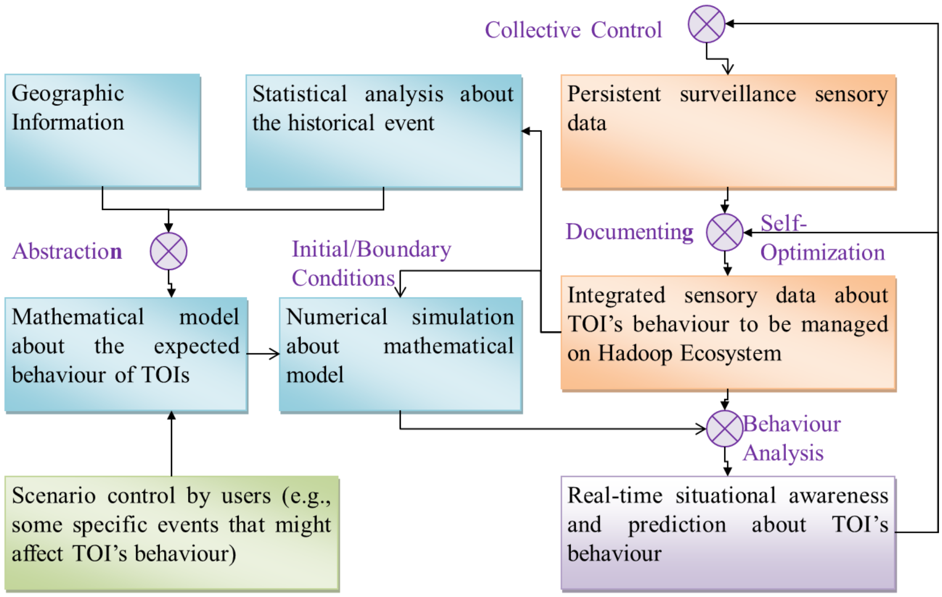

Fig. 4 A generic flow-chart about sensor-oriented information analysis system.

Processing of persistent surveillance video data includes the following operations: (1) acquisition of video data; (2) segmentation, which extracts pixels of target-of-interests (TOI) from background; (3) isolation of TOIs out of noise or other moving objects; (4) translation of optical behavior features (i.e., the velocity and position of moving TOIs within the sensor coordinate system) of detected pedestrians into their actual geographical features (i.e., the velocity and position of moving targets within the geographic coordinate system); (5) documentation, which posts the output in a format suitable for post-processing and includes position, velocity, and track. The implementation of both threads is highly computation and storage-intensive.

5. Data Analytics Based on Mathematical Model

In the context of sensor-oriented analysis, data analytics of sensory data aims to disclose specific behavior of target-of-interest (TOI) out of sensory data [9][10][16]. For example, it is a significant task to detect those speeding or wrong-way vehicle out of surveillance video on the road; therefore a description about the expected (or normal) traffic flow is needed so that those abnormal vehicles can be identified. The expected behavior (i.e., normal behavior) of TOI is commonly modeled using macroscopic or microscopic method. Microscopic method, which is also called agent-based method, provides a detailed formulation about the behavior of TOI while suffers from inhibitive computational cost and accumulated numerical error. Macroscopic method generally uses time- and space-dependent partial differential equations (PDE) to formulate the expected behavior of TOI. In this paper, microscopic and macroscopic methods are used to simulate the aggregation of carp [26-33] and vehicle traffic flow [34-36] respectively.

5.1. Aggregation of Carp

Since being introduced to the U.S. in the 1970 for the purpose of weed and parasite controlling in aquatic farms [37-39], the Asian carp (including bighead carp, the black carp, the grass carp, and the silver carp) has gradually established breeding populations in Mississippi River region [39-41]. Asian carps are causing serious damage to the area' fresh-water ecosystem [38][40][42]. In order to provide constructive information to control the populations of Asian carps [26][43], the addressed SOIS-Hadoop is customized and employed to predict the collective behaviour of Asian carps (Figure 5(a)).

In this work, an agent-based mathematics model (a microscopic method) is presented to formulate the aggregation of Asian carps. Based on the statistical analysis about empirical sensory data, the pair-wise interaction U_{ij} is defined using modified Van der Waals forces [44], where the corresponding potential function U_{ij} between carp-i and carp-j is defined by the Formula (1).

$$U_{ij} = \begin{cases} \varepsilon\left[\left(\frac{\sigma}{\|r_{ij}\|}\right)^{M} - \left(\frac{\sigma}{\|r_{ij}\|}\right)^{N}\right] & (\| r_{ij} \| < R_s) \\ \varepsilon\left[\left(\frac{\sigma}{R_s}\right)^{M} - \left(\frac{\sigma}{R_s}\right)^{N}\right] & (R_s \le \| r_{ij} \| \le R_h) \\ \varepsilon\left[\left(\frac{\sigma}{\|r_{ij}\| - R_h + R_s}\right)^{M} - \left(\frac{\sigma}{\|r_{ij}\| - R_h + R_s}\right)^{N}\right] & (R_h < \| r_{ij} \| < R_k) \end{cases} \tag{1}$$

where $r_{ij} = X_i - X_j$, $M > N$, and $\sigma = R_s\left(\frac{N}{M}\right)^{\frac{1}{M-N}}$ (such that $\nabla U_{ij}(\|r_{ij}\|) = 0$ when $R_S \le r_{ij} \le R_h$). From Formula (1), it follows that the resulting force function is:

$$\begin{aligned} f_{ij} &= -\frac{\partial U_{ij}}{\partial \| r_{ij} \|} \\ &= \begin{cases} -\frac{\varepsilon}{\sigma}\left[M\left(\frac{\sigma}{\|r_{ij}\|}\right)^{M+1} - N\left(\frac{\sigma}{\|r_{ij}\|}\right)^{N+1}\right] & (\| r_{ij} \| < R_s) \\ 0 & (R_s \le \| r_{ij} \| \le R_h) \\ -\frac{\varepsilon}{\sigma}\left[M\left(\frac{\sigma}{\|r_{ij}\| - R_h + R_s}\right)^{M+1} - N\left(\frac{\sigma}{\|r_{ij}\| - R_h + R_s}\right)^{N+1}\right] & (R_h < \| r_{ij} \| < R_k) \end{cases} \end{aligned} \tag{2}$$

Compared to alternative models such as [33][45][46], the addressed model can efficiently formulate the "aggregation" of carp school.

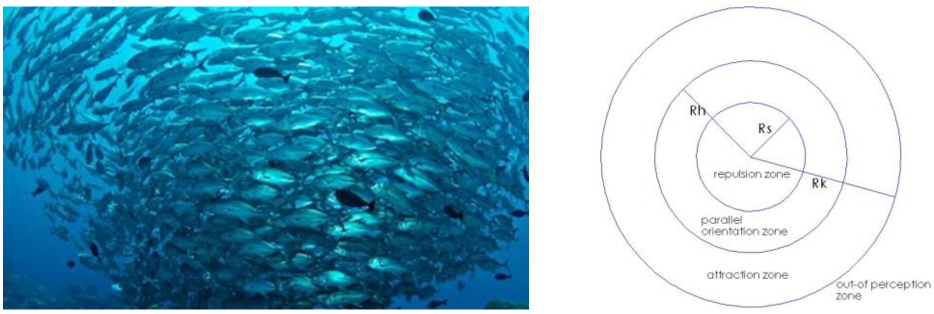

Fig. 5 (a) aggregation of carp; (b) Interaction zones between neighboring carp.

R_s, R_h, and R_k are illustrated in Fig. 5(b), $||r_{ij}||$ indicates the distance between two neighboring carps. ε is constant coefficient derived from empirical data. It should be remarked that the moving orientation, water flow velocity and blind zone is not considered in the formulation of Formula (1).

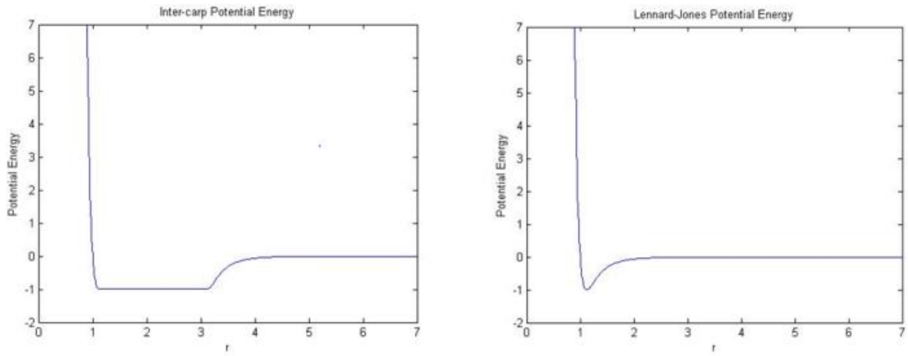

Fig. 6 (a) inter-carp potential energy; (b) inter-carp force (M = 12; N = 6)

Fig. 6(a) shows the potential energy incurred by the pair-wise interaction between two neighboring carps. Fig. 6(b) shows the resulting inter-carp force. It can be observed that, inter-carp potential energy has a stable zone (or parallel zone), within which the inter-carp potential energy is basically constant so that the neighboring carps can cruise without influencing each other.

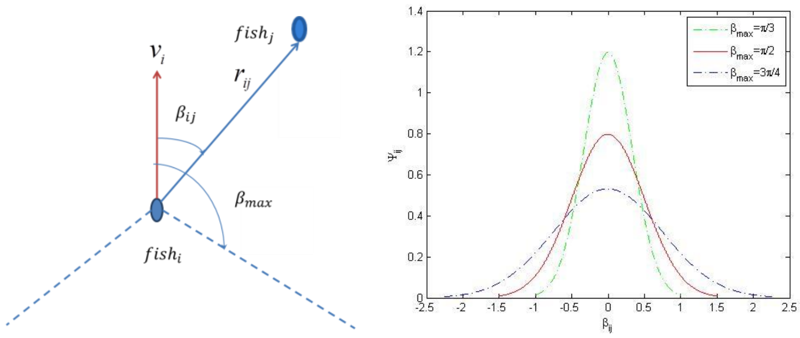

Fig. 7 (a) Interaction between neighboring carps;
(b) Ψ_{ij} value with variant $\beta_{\max}$ (denoting the visible zone)

According to ichthyology [31-32], the interaction between carps is supposed to be corresponding to blind-zone (Figures 5(b) and 7(a)). As illustrated in the Figure 7(a), ν_i is the velocity of i-th carp. $\beta_{\max}$ is the maximal perceptible angle, obviously $0 \le \beta_{\max} \le \pi$. β_{ij} indicates the angle between ν_i and r_{ij}, it is defined by the following formula:

$$\beta_{ij} = \arccos\left(\frac{\nu_i \circ r_{ij}}{||\nu_i||\ ||r_{ij}||}\right) \tag{3}$$

Given the blind-angle $\beta_{\max}$, the inter-carp potential is defined as:

$$U_{ij}^{*} = \Psi_{ij} U_{ij} \tag{4}$$

Where

$$\Psi_{ij} = \Psi(\beta_{ij}) = \frac{\sqrt{\pi}}{\beta_{\max}\sqrt{2}} e^{\frac{-\pi^2 \beta_{ij}^2}{2\beta_{\max}^2}} \tag{5}$$

As a consequence, the inter-carp interaction force is determined by the following formula:

$$F_{ij}^{*} = \frac{\partial}{\partial \| r_{ij} \|} U_{ij}^{*} = \Psi_{ij} f_{ij} \tag{6}$$

where f_{ij} is defined in Equation (2).

Based on the above mathematical model for carp schooling, the future status of carp school can be predicted according to the currently observed sensory data using agent-based mathematics model addressed above. Furthermore, based on the preliminary simulation results, the motivations of fish aggregation, such as foraging advantages, reproductive advantages, predator avoidance, or hydrodynamic efficiency, can be disclosed.

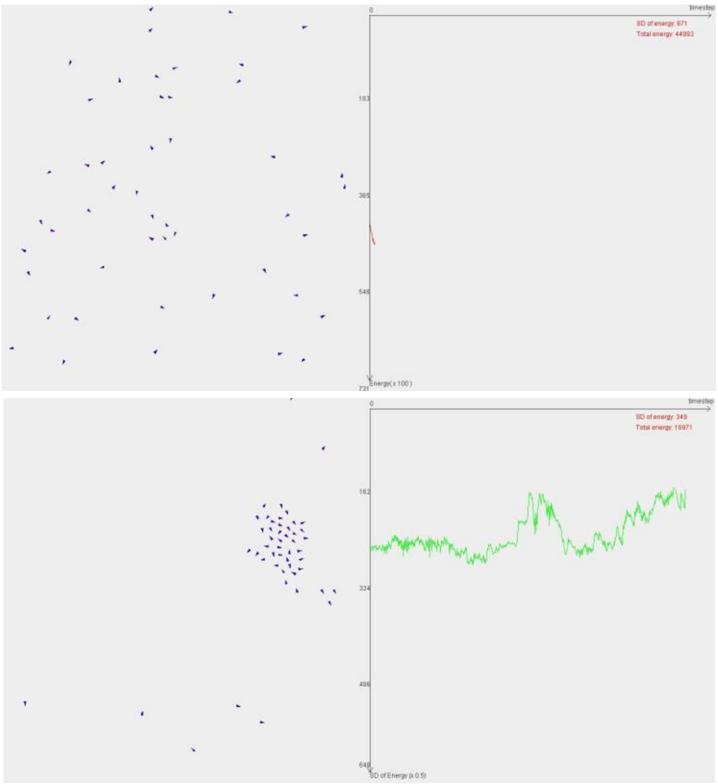

Fig. 8 Snapshots about the simulation of carp aggregation and corresponding standard-deviation of kinetic energy (the size of fish school is 50): (a) initial stage; (b) aggregation stage

Figure 8 demonstrated the aggregation process of a fish school of size 50. It is illustrated that carp gradually gather due to the pairwise interaction between neighboring carp; in addition, standard-derivation of kinetic energy of carp can be used to measure the aggregation status of carp school, namely a carp school in aggregation has smaller standard deviation of kinetic energy.

5.2. Vehicle traffic analysis

Traffic flow analysis plays a significant role in civil engineering, transportation management, and homeland security [34]. Due to the influence of rapid urbanization and modern industrialization, traffic congestion has become an intolerable issue in today's world.

Modeling and simulation of traffic flow provides an efficient way to understand traffic congestion and disclose corresponding remedy. Mathematical models for traffic flow are categorized into microscopic (or agent-based) and macroscopic strategies. Macroscopic models study traffic from an average (or continuum) perspective, while microscopic models study the motion of individual vehicles.

Macroscopic model uses temporal and spatial-dependent partial differential equations (generally hyperbolic partial differential equations.) to formulate the expected traffic flow. Representative macroscopic models for traffic flow are Lighthill-Whitham-Richard model [35], Aw-Rascle Model [47], and Zhang Model [36]. None of the above models can efficiently and accurately those formulate complicated road scenario such as nozzle, merging, diverging, and roundabout, etc.

Different from above models, the proposed work defines the governing equations for traffic flow using the following partial differential equations:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho V) = 0 \tag{7.1}$$

$$\rho(X,t) = \frac{r}{V_{\max} A(X) \log A(X)} \tag{7.2}$$

$$\rho \frac{DV}{Dt} = -\nabla P + \mu \nabla^2 V + \rho g \tag{7.3}$$

Where $\rho(X,t)$ is the number of vehicles over unit length, $V(X,t)$ is the expected velocity of the vehicle, $V_{\max}$ is the speed limit, $A(X)$ is the cross-section width (or bandwidth) of the road. Equation (7.1) is derived from conservation of mass. Equation (7.2) ensures that traffic flow slows down up at nozzle and keeps constant speed at fork (illustrated in Figure 9).

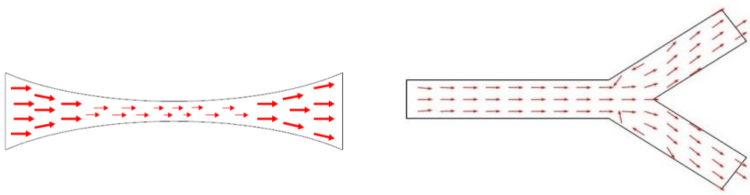

Fig. 9 Traffic flow through (a) nozzle; (b) fork

As illustrated in Figure 10(a), this work acquires the citywide traffic status using electro-optical sensor array mounted on Unmanned Aerial Vehicle (UAV). Figure 10(b) shows the expected traffic velocity field resulted from the solution of governing equations. The boundary conditions and the coefficient equations for governing equations are obtained according to empirical traffic data. Using the expected traffic flow as reference, the observed vehicles can be measured and evaluated.

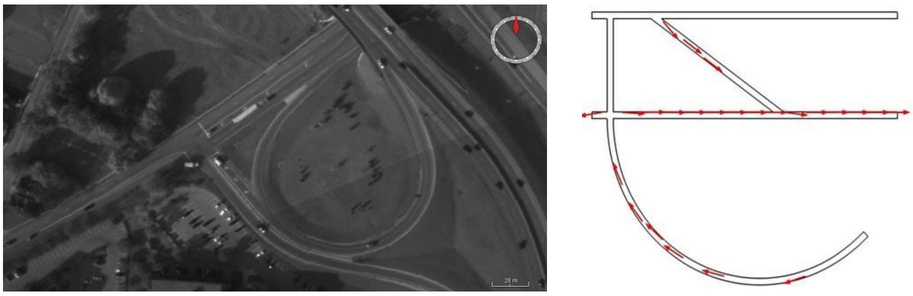

Fig. 10 (a) traffic status acquired using UAV-mounted optical-electro sensor array; (b) expected traffic flow derived from empirical data and Equation (6).

6. Macro-cell Strategy

Real-world problem generally involves a large scene such as a metropolitan city, or a huge lake. As a result, a SOIS-Hadoop system should be scalable so as to solve the large-scale problems.

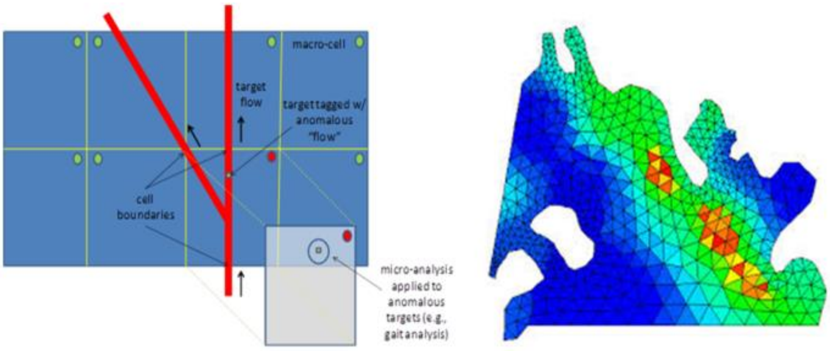

Fig. 11 Two partition strategies: (a) Euler formation of a transportation network; (b) Lagrange formulation of a lake.

In this work, a macro-cell strategy, which partitions the global physics domain (or scene) into multiple overlapping/non-overlapping element (cell) and then manipulates them independently [9][10][16][48][49], will be employed to enhance the capability of the SOIS-Hadoop framework to handle large-scale problems. As illustrated in Fig. 8, the physics domain (or scene) of interest can be discretized using Euler formulation or Lagrange formulation [49]. Inter-cell communications only occur between neighboring cells and they are only triggered while somewhat anomalous crowd behavior is observed and detected. Cell of particular interest will be particularly analyzed using modeling and simulation strategy (which is relatively computationally costly).

Table 1 lists sample features value about macro-cell-oriented carp aggregation analysis [46] methods. Through sufficient training, the addressed system can accurately employ the known cellular features to predict the likelihood of aggregation occurrence through appropriate machine learning methods [50] such as logistic regression, neural network, Hidden Markov Method (HMM), and Bayesian learning, etc.[50].

Table 1 Cell-by-cell analytics of carp aggregation (the lake is divided into 1000 cells).

Cell ID	Fish density	Total Kinetic Energy	STD (kinetic -energy)	Entropy	Aggregation occurs?
1	25.6	77	10.2	10	Yes
2	12.7	89	30.98	40	No
3	7.9	101	105	90	-
…	..			…	…
10	14	25	133	30	-

7. Conclusion

A pilot SOIS-Hadoop system has been set-up and applied in a variety of real-world problems such as the prediction of the aggregation of carp [16] and vehicle traffic analysis [6][9][24]. Some preliminary while promising outcomes has achieved.

In the near future, we intend to make progress in the following directions: (1) Broaden the application of the proposed sensor-oriented information analysis system such as the simulation about the spread of epidemics diseases [16], anomalous pedestrian detection [8][15], and structural health monitoring,, etc.; (2) Develop scalable numerical methods in the mathematical modeling of sensor oriented information analysis system: time integration method for the solution of governing equation, domain decomposition method in the finite element method, and polynomial preconditioning, etc.; (3) Optimize the exploitation of sensory data using dimensionality reduction (e.g. such as PCA) [50]; (4) Optimize the cooperative control of sensor asset so as to obtain the optimal observation and high energy efficiency; (5) Employ more advanced and accurate mathematical model to formulate the expected behavior about TOI. For example, stochastic analysis can be introduced to formulate uncertainty of SOIS-Hadoop framework; and multi-scale modeling can be used to a seamlessly merge microscopic and macroscopic description about TOI.

Acknowledgement: *This work is jointly sponsored by the National Science Foundation (NSF) with proposal number 1240734 ("A Design Proposal for the Center of Cyber Sensor Networks for Human and Environmental Applications") and 1111542 ("RI: Large: Collaborative Research: A Robotic Network for Locating and Removing Invasive Carp from Inland Lakes"). The authors would like to thank Dr. Kimberly Kendrick from University of Nevada -- Las Vegas (Nevada, USA), Dr. Xiaofang Wei from Central State University (Ohio, USA), Mr. Darrell Barker and Ms. Olga Mendoza-Schrock of the Sensors Directorate in Air Force Research Laboratory (Ohio, USA) for their support and guidance of this work.*

REFERENCES

[1] O. Bott, M. Marschollek, K.-H. Wolf, and R. Haux, "Towards new scopes: sensor-enhanced regional health information systems-part 1: architectural challenges", *Meth. Inform. Med.,* Vol. 46, pp. 476-483, 2007.
[2] J. V. C. Schneider, Information Systems Today: Managing in the Digital World. Prentice Hall, 2015.
[3] F. Zhao, J. Shin, and J. Reich, "Information-driven dynamic sensor collaboration", *IEEE Signal Process. Mag.,* Vol. 19, pp. 61-72, 2002.
[4] W. M. Ulrich, Legacy systems: transformation strategies. Prentice Hall, 2002.
[5] S. Fernandes, Y. Liang, S. Sritharan, X. Wei, and R. Kandiah, "Real Time Detection of Improvised Explosive Devices using Hyperspectral Image Analysis", In Proceedings of the 2010 IEEE National Aerospace and Electronics Conference (NAECON 2010). 2010.
[6] S. Fernandes and Y. Liang, "Chipping and segmentation of target of interest from low-resolution electro-optical data", In Proceedings of the SPIE Defense, Security, and Sensing. 2013, pp. 87440R-87440R-8.
[7] K. Grolinger, W. A. Higashino, A. Tiwari, and M. A. Capretz, "Data management in cloud environments: NoSQL and NewSQL data stores", *J. Cloud. Comput. Adv. Syst. Appl.,* vol. 2, p. 22, 2013.
[8] Y. Liang, W. Melvinb, S. Fernandesa, M. Hendersona, S. I. Sritharanc, and D. Barkerd, "A Crowd Motion Analysis Framework Based on Analog Heat-Transfer Model", *American Journal of Science and Engineering,* Vol. 2, pp. 33-43, 2013.
[9] Y. Liang, M. Henderson, S. Fernandes, and J. Sanderson, "Vehicle tracking and analysis within a city", In Proceedings of the SPIE Defense, Security, and Sensing. 2013, pp. 87510F-87510F-15.
[10] Y. Liang, M. Szularz, and L. T. Yang, "Finite-element-wise domain decomposition iterative solvers with polynomial preconditioning", *Math. Comput. Model.,* vol. 58, pp. 421-437, 2013.
[11] A. S. Foundation. (2014). *Hadoop Releases*. Available: http://www.apache.org/
[12] Y. Liang and C. Wu, "A Sensor-Oriented Information System Based on Hadoop Cluster", In Proceedings on the International Conference on Internet Computing (ICOMP). 2014, p. 1.
[13] Y. Liang and C. Wu, "An agent-based mathematical model about carp aggregation", In Proceedings of the SPIE Sensing Technology+ Applications. 2015, pp. 94860Q-94860Q-11.
[14] R. C. Taylor, "An overview of the Hadoop/MapReduce/HBase framework and its current applications in bioinformatics", *BMC Bioinformatics,* vol. 11, p. S1, 2010.
[15] Y. Liang, W. Melvin, S. I. Sritharan, S. Fernandes, and D. Barker, "CMA-HT: a crowd motion analysis framework based on heat-transfer analog model", In Proceedings of the SPIE Defense, Security, and Sensing. 2012, pp. 84020J-84020J-13.
[16] Y. Liang, Z. Shi, S. I. Sritharan, and H. Wan, "Simulation of the Spread of Epidemic Disease Using Persistent Surveillance Data", In proceeding of COMSOL 2010, Boston, 2010.
[17] K. Langendoen and N. Reijers, "Distributed localization in wireless sensor networks: a quantitative comparison", *Comput. Netw.,* Vol. 43, pp. 499-518, 2003.
[18] H. Qi and J. B. Moore, "Direct Kalman filtering approach for GPS/INS integration", *IEEE Trans. Aerosp. Electron. Syst.,* Vol. 38, pp. 687-693, 2002.
[19] G. A. Bekey, Autonomous robots: from biological inspiration to implementation and control. MIT press, 2005.
[20] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey", *Ad hoc networks,* Vol. 7, pp. 537-568, 2009.
[21] T. S. Rappaport, Wireless communications: principles and practice. Vol. 2: prentice hall, 2002.
[22] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: a survey", *IEEE Wireless Commun.,* Vol. 11, pp. 6-28, 2004.
[23] Y. Liang, J. Weston, and M. Szularz, "Generalized least-squares polynomial preconditioners for symmetric indefinite linear equations", *Parallel. Comput.,* Vol. 28, pp. 323-341, 2002.
[24] J. Sanderson and Y. Liang, "No-reference image quality measurement for low-resolution images", in SPIE Defense, Security, and Sensing. 2013, pp. 874404-874404-17.
[25] W. Ren and R. W. Beard, Distributed consensus in multi-vehicle cooperative control. Springer, 2008.
[26] S. Camazine, *Self-organization in biological systems*. Princeton University Press, 2003.
[27] I. D. Couzin, J. Krause, R. James, G. D. Ruxton, and N. R. Franks, "Collective memory and spatial sorting in animal groups", *J. Theor. Biol.,* Vol. 218, pp. 1-11, 2002.
[28] A. Deutsch and S. Dormann, Cellular automaton modeling of biological pattern formation. 2005.
[29] A. Huth and C. Wissel, "The simulation of the movement of fish schools", *J. Theor. Biol.,* Vol. 156, pp. 365-385, 1992.
[30] P. B. Johnsen and A. D. Hasler, "Winter aggregations of carp (Cyprinus carpio) as revealed by ultrasonic tracking", *Trans. Am. Fish. Soc.,* Vol. 106, pp. 556-559, 1977.
[31] R. Jullien and R. Botet, *Aggregation and fractal aggregates*. World Scientific Pub Co Inc, 1987.
[32] S. Stöcker, "Models for tuna school formation", *Math. Biosci.,* Vol. 156, pp. 167-190, 1999.

[33] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles", *Phys. Rev. Lett.,* Vol. 75, p. 1226, 1995.

[34] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, "Dynamical model of traffic congestion and numerical simulation", *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.,* Vol. 51, p. 1035, 1995.

[35] M. J. Lighthill and G. B. Whitham, "On kinematic waves. II. A theory of traffic flow on long crowded roads", In Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. 1955, pp. 317-345.

[36] H. M. Zhang, "A mathematical theory of traffic hysteresis", *Transport. Res. B-Meth.,* Vol. 33, pp. 1-23, 1999.

[37] A. C. R. C. Committee, "Asian carp control strategy framework", 2013.

[38] E. H. Buck, H. F. Upton, C. V. Stern, and J. E. Nicols, "Asian carp and the Great Lakes region", 2010.

[39] J. Rosenfeld, "Assessing the habitat requirements of stream fishes: an overview and evaluation of different approaches", *Trans. Am. Fish. Soc.,* Vol. 132, pp. 953-968, 2003.

[40] R. Naylor, S. Williams, and D. Strong, "Aquaculture-a gateway for exotic species", *Science (Wash.),* Vol. 294, pp. 1655-1656, 2001.

[41] R. Goldburg, M. S. Elliott, R. Naylor, and P. O. Commission, *Marine aquaculture in the United States: environmental impacts and policy options*. Pew Oceans Commission, 2001.

[42] T. M. Koel, K. S. Irons, and E. N. Ratcliff, "Asian carp invasion of the upper Mississippi River system", US Department of the Interior, US Geological Survey, Upper Midwest Environmental Sciences Center, 2000.

[43] P. Bajer, C. Chizinski, and P. Sorensen, "Using the Judas technique to locate and remove wintertime aggregations of invasive common carp", *Fish. Manage. Ecol.,* Vol. 18, pp. 497-505, 2011.

[44] A. R. Leach, Molecular modelling: principles and applications. Prentice Hall, 2001.

[45] A. Czirók, M. Vicsek, and T. Vicsek, "Collective motion of organisms in three dimensions", *Physica. A.,* Vol. 264, pp. 299-304, 1999.

[46] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model", in ACM Siggraph Computer Graphics. 1987, pp. 25-34.

[47] A. Aw and M. Rascle, "Resurrection of" second order" models of traffic flow", *SIAM journal on applied mathematics,* Vol. 60, pp. 916-938, 2000.

[48] Y. Liang, The Use of Parallel Polynomial Preconditioners: In the Solution of Systems of Linear Equations. LAP LAMBERT Academic Publishing, 2013.

[49] A. Toselli and O. Widlund, *Domain decomposition methods: algorithms and theory.* Vol. 3: Springer, 2005.

[50] S. Theodoridis, Machine Learning: A Bayesian and Optimization Perspective. Academic Press, 2015.