Surrogate Mobile Sensing

Dong Wang[1], Tarek Abdelzaher[2], Lance Kaplan[3]
[1]Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN 46556
[2]Department of Computer Science, University of Illinois at Urbana Champaign, Urbana, IL 61801
[3]Networked Sensing and Fusion Branch, US Army Research Labs, Adelphi, MD 20783

***Abstract*—The proliferation of smart phones with sensing capabilities motivates exploring the applicability limits of (phone-based) mobile sensing. While a phone can directly measure variables such as location, acceleration, and orientation, other interesting quantities that one may want to measure have higher-level semantics that a phone does not directly recognize. For example, one might want to map parking lots that are free after hours, or restaurants that are popular after midnight. How can we measure such higher-level logical quantities using sensors on phones? Techniques that address this question fall in the broad area of surrogate sensing, defined as inferring high-level logical quantities by measuring weaker surrogates. The surrogates in question are variables that can be sensed using a phone, but that are only weakly related to the original high-level logical quantities one is really after. The key challenge is to exploit appropriate aggregation techniques that leverage the availability of large numbers of phones to overcome the poor quality of individual surrogates. Recently, significant advances were made on understanding the quality limits of surrogate sensing. This article overviews the main ideas and insights underlying these advances.**

I. Introduction

Mobile crowd sensing [1] is a sensing paradigm that exploits the ubiquity of mobile devices. Recent years have seen a significant rise in mobile crowd sensing applications [2]–[4] and frameworks [5], where individuals, or devices in their possession, collect measurements that allow mapping quantities of interest at scale. The power of this paradigm depends on the sensing capabilities of mobile devices and what can be inferred from their measurements using appropriate fusion techniques.

Measurements such as location, speed, acceleration, and orientation can be easily made on a mobile device. Other quantities of interest, however, may have more complex semantics, not directly understood by mobile sensors. For example, one might want to find "a free parking lot" or "a good restaurant", but phone sensors do not explicitly measure these quantities. One way to address the problem is to involve people as "sensors" to observe and report such finds [6]. In this article we explore an alternative, called *surrogate sensing*, where common mobile device sensors are re-purposed (in software) to do the job on their own, utilizing only their measurements and appropriate data aggregation techniques. An analytic approach is described that is intended to offer a "force multiplier" significantly extending the applicability of mobile sensors.

Surrogate sensing refers to *transparent* use of commonly available sensors (e.g., on mobile phones) as surrogates for measuring more complex logical quantities. These surrogates are individually less accurate (compared to direct human observations). Nevertheless, putting many surrogate measurements together can, under certain conditions, reliably reproduce the high-level logical quantities of interest. The main challenge is to come up with algorithms for aggregation that produce good quality results from weak surrogates and quantify the error bounds on the results produced.

To give an example, consider the problem of identifying parking lots that are free of charge after 5pm in a given town. Short of asking for people's cooperation (e.g., by explicit tagging), this sensing problem is hard to accomplish on a phone. An approximate solution could be to write a "stop detection" application. Whenever a driver stops and gets out of their car after 5pm, the application automatically tags this location as a free-of-charge parking lot (essentially making the assumption that, given the choice, most people would usually prefer not to pay for parking). In this case, sensing the "stop" event after 5pm, which can be done on a phone, is a surrogate for sensing a free-of-charge parking lot. Clearly, this surrogate may result in many misclassifications. There are many reasons why individuals might park at pay parking as well. Hence, one question is: can one reliably detect free-of-charge parking using such an unreliable surrogate? Recent results show that this and similar questions can often be answered in the affirmative as long as a sufficiently large number of measurements is present. Moreover, error bounds have been derived that describe the quality of surrogate sensing results.

Informally, the solution consists of two inter-dependent parts. Namely, (i) identify "good sources", and (ii) give these sources more weight when conflicting sensing results occur. For example, some individuals are more price-conscious than others. Identifying these individuals and monitoring where they park offers a better surrogate for finding free parking lots. The approach enables (sufficiently large sets of) phones to transparently and reliably measure logical quantities whose semantics the individual phones cannot directly recognize. Note that, a key ingredient of this genre of sensing is to exploit the behavior of individual sources, zooming in on those that act as good surrogates for the quantities we want to measure. Hence, a *surrogate* is a sensory measurement of a given *source*.

The above discussion demonstrates that surrogate sensing is challenging not only because we need to figure out the values of the logical quantities that a phone cannot directly measure, but also because we may not know upfront how good or bad the used surrogates are. Hence, we must jointly (i) recognize the quality of sources in offering good surrogate indicators, and (ii) infer the values of the sought logical quantities.

Recent work casts surrogate sensing as a maximum likelihood estimation problem [7]. The general idea is to exploit mutual consistency relations among data items, as well as consistency with physical constraints. Together, these relations and constraints shape the likelihood function that quantifies the odds of the observations at hand. We then maximize the resulting likelihood function with respect to hypotheses on the correctness of individual observations as well as hypotheses on the quality of individual surrogates. The maximum likelihood estimate thus obtained offers a best guess of both the logical quantities in question and the quality of used surrogates.

Importantly, the error variance of (unbiased) maximum likelihood estimators is well-known. By casting the surrogate sensing problem as one of maximum-likelihood estimation, it becomes possible to compute the error variance and hence offer confidence intervals in estimation results. These intervals allow one to understand how good these results are. Hence, not only can we allow phones to measure high-level logical quantities using weak surrogates, but also we can offer probabilistic assurances in result quality. The rest of this article elaborates on the techniques used and their performance.

II. Binary Surrogate Sensing

The simplest surrogate sensing problem is one that involves measurement of binary logical variables. While a binary model might sound restrictive at first, it may actually capture arbitrary logical facts and their negation. For example, all geo-tagging applications can be thought of as estimation of binary variables representing the presence or absence of certain entities at different locations. A running example we use in this article is to locate stop signs in a city. In this case, each intersection can be associated with a set of binary variables, each indicating the presence or absence of a stop sign at the corresponding road at the intersection.

Since most vehicles stop briefly at stop signs, a surrogate might be to identify those instances where a vehicle stopped for a short period of time (say, between 2 and 10 seconds), as an indication of the existence of a stop sign. Clearly, there may be other reasons for a vehicle to stop, but the beauty of surrogate sensing lies in that the surrogate does not have to be accurate. Hence, a simple smart-phone app can be written that gets activated when the phone senses that it is in a car (say, from the profile of acceleration measurements). The app would then monitor the car's GPS location, identify instances when the location remains unchanged for the indicated short period of time, and if so, make the binary claim that a stop sign has been encountered at the corresponding location. Note that, some drivers are more consistent at fully stopping at stop signs while others come to a rolling stop that does not trigger a claim. Also, drivers may stop for other reasons (e.g., cabs may pick up passengers). Hence, when using a short stop as a surrogate for a stop sign, the quality of that surrogate depends on the driver.

A server receives the binary measurements from the community of phones. The question is how to determine, given only the measurements sent and without knowing the quality of surrogates upfront, which of the reported observations are true and which are not. In other words, which are real stop signs and which are not.

A trivial way of solving this problem is by "believing" only those observations that are reported by a sufficient number of sources. This scheme is called *voting*. The problem with voting schemes is that they do not take source reliability into account. Instead, it is better to identify good surrogates and give them a higher weight in voting.

The problem of jointly uncovering source reliability and observation correctness can be traced back to Google's PageRank [8]. PageRank iteratively ranks sources on the Web, by considering their linkages. Extensions of PageRank, known as fact-finders, iteratively compute the quality of sources and their claims. Specifically, they estimate the credibility of claims from the quality of sources that make them, then refine the quality of sources based on the credibility of their claims. Several algorithms exist that feature modifications of the above basic scheme [9]. Surrogate sensing builds on the above ideas by considering different sources (the surrogates), each indirectly indicating values of some logical quantities (the claims), then solving for the reliability of sources and claims.

A. An Optimal Solution

An optimal solution to the surrogate sensing problem is to design a maximum likelihood estimator that considers a group of surrogates who indicate a set of logical binary variables. One may consider the default value of each variable to be zero (e.g., the default is "no stop sign"). Hence, a surrogate needs to indicate only when a variable deviates from the default (i.e., is 1). It is useful to group the inputs into an input matrix, whose dimensions are the number of surrogates (sources) and the number of binary variables. The cells in the matrix are binary values that specify which surrogate indicated which variable. These indications are reported to a server as binary *claims*. We do not, in general, know whether any individual claim is correct or not.

We describe the quality of each surrogate by the probability of true positives (probability that it reports a binary variable correctly) and the probability of false positives (probability that it makes a bad claim). The maximum likelihood estimator finds the values of these probabilities that maximize the likelihood of received inputs (i.e., the values reported in the input matrix). We solve this problem using a standard expectation maximization algorithm.

B. Reliability Assurances

The expectation maximization approach, described above, has an interesting property. Namely, under very general conditions, it enables computing *confidence* in results [10]. Hence, not only do we arrive at conclusions regarding (i) correctness of individual observations and (ii) quality of individual surrogates, but also we obtain a level of confidence in each conclusion. Intuitively, when the algorithm estimates that a source offers a reliable surrogate 50% of the time, there is a big difference in confidence depending on whether this conclusion is based on two measurements (one estimated to be right and one estimated to be wrong) or 100 measurements (50 estimated to be right and 50 estimated to be wrong). The ability to estimate confidence in results is one of the strong advantages of the statistical approach.

In recent work [11], simulation results were reported for three different system scales: small, medium and large, featuring 100, 1000, and 10000 sources, respectively. Each source represented a surrogate of different quality. The EM algorithm was executed to assess the quality of different surrogates.

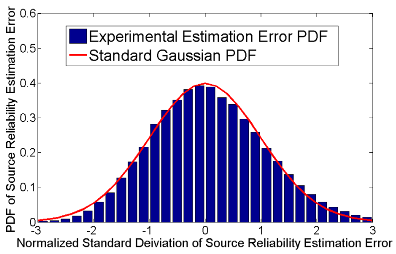

Fig. 1. Normalized Source Reliability Estimation Error PDF

Figure 1 depicts the error incurred by the maximum-likelihood estimator is assessing the quality of surrogates (where the horizontal axis shows the error, and the vertical shows the percentage of time it occurs). It follows a normal bell-shaped curve (i.e., the Gaussian distribution). The curve allows one to assess how good the estimator is. Importantly, it was shown that this curve can also be derived mathematically [11]. Hence, *analytical predictions* can be made about quality, such as: what value of error is not exceeded by 95% of the estimates? Results, reported in prior work [11], show that the difference between analytical predictions and empirical measurements was less than 1%.

The above results are significant in that they demonstrate feasibility of assessing, with statistical confidence, the quality of different surrogates given only the information on which surrogate indicated which values. The estimate is derived with neither prior knowledge of the reliability of surrogates nor means for directly verifying the correctness of their measurements. The attainment of assurances in the absence of requirements for prior knowledge enables surrogate sensing to offer practical correctness guarantees.

A careful read of the above results might lead to the question of whether some form of circular reasoning is involved in these analytic foundations. After all, how can one assess quality of both surrogates and logical variables they measure at the same time when one depends on the other? Indeed, there are cases, where the iterative algorithm becomes a "self-fulfilling prophecy". For example, if one starts with the assumption that all surrogates are correct and all variables are true, the maximum likelihood estimation remains trapped at this initial condition because it is self-consistent. The same is true if one starts with the assumption that all surrogates are wrong and all variables are false. In general, there is symmetry in the solution space; by changing all "true" values to "false" and all probabilities of "correctness" to probability of "error" the consistency of a solution is not compromised. In other words, there is an inherent ambiguity in data interpretation; either some sources are good and their observations are correct, or these sources are bad and the negation of their observations is correct. For the approach to converge reliably to the right conclusions, there needs to be some disambiguation.

Informally, the disambiguation that causes the iterations to converge in the right direction lies in an implicit assumption of coherence of truth (at least most of the time). For example, if locations of stop signs are reported, random stops at other locations should be less systematic. Moreover, the pool of drivers who stop at those other locations should be somewhat independent from those who always stop at stop signs. Intuitively, favoring the likelihood that consistent observations are true, the iterations boost the (credibility of) sources involved in those observations and accordingly the other observations by those sources. Other sources and their observations receive a lower credibility. Hence, a random initial assignment usually causes iterations to converge to a unique non-trivial solution that is independent of the initial assignment and that is shown to outperform other baselines in its correspondence with ground truth [7]. An open problem remains to identify exactly the conditions under which convergence to a unique optimal solution is guaranteed.

C. Non-independent Observations

The results of maximum likelihood estimation, presented above, can be further improved by considering a number of constraints [12]. Informally, the constrained optimization problem searches for the best estimate of the unknown logical quantities within a smaller space than its unconstrained counterpart. Hence, the accuracy of results is improved. Two types of constraints are suggested:

- *Source/measurement co-location constraints:* These constraints state that a source (i.e., a smart phone) can only measure variables at locations that it visited. Hence, if we know the locations visited by the source, we can better interpret silence (i.e., lack of claims). For example, if a given phone did not indicate the presence of a stop sign at a given intersection, should that silence decrease our confidence that a stop sign is present there, or not? The answer depends on whether that phone visited the intersection (in a car) or not. Therefore, having access to source location traces can improve the quality of the maximum likelihood estimation.
- *Measurement correlation constraints:* The values of logical variables we wish to determine may be correlated, which can be expressed by a joint probability distribution. For example, on a large road, no stop signs are usually present at intersections. Hence, the existence of stop signs at intersections of the same road may be correlated. Taking these correlations into account in the maximum likelihood estimation problem allows the estimator to better assess actual values of the underlying logical variables.

The aforementioned constraints have been integrated with the basic maximum-likelihood formulation [12]. Specifically, to accommodate source/measurement co-location constraints, the algorithm does not penalize sources for not reporting measurements that they did not have an opportunity to observe. It also does not take it against reported measurements if they are not corroborated by sources who never had an opportunity to observe them. Similarly, to accommodate measurement correlation constraints, the algorithm takes into account the joint probability distribution of correlated variables when assessing the likelihood of correctness of reported observations of these variables. Evaluation results, reported later, show that these optimizations appreciably improve accuracy.

III. Multi-valued Sensing

The binary surrogate sensing problem described in the previous section can be viewed as a two-way classification problem, where surrogates are used to categorize logical quantities into one of two classes. For example, roads at intersections are categorized into the classes "with stop sign" or "without". This is a special case of the more general multi-label classification problem, which can model multi-valued surrogate sensing.

Using the above insight, it is possible to extend the formulation of the maximum likelihood estimator to account for an arbitrary number of classes. In this case, surrogates will indicate one of multiple values. For example, the surrogate for detecting stop signs can be generalized to detecting other traffic regulators. In the generalized app, a short stop (say, 2-10 seconds) can still indicate a stop sign. A long stop (say, 15-90 seconds) can indicate a traffic light. No stop, when consistently indicated, means the absence of a regulator. Hence, a three-way classifier can be used to bin each intersecting street into one of three categories depending on the presence and type of the involved traffic regulator; the three categories being *traffic-light*, *stop-sign*, and *none*.

As before, different surrogates can have different levels of quality. For example, cars whose drivers travel primarily at rush hour may see delays larger than 90 seconds at traffic lights and larger than 15 seconds at stop signs. These larger delays will therefore increase false negatives (e.g., traffic lights not recognized by the surrogate) and misclassifications (e.g., stop signs mistaken for traffic lights). In contrast, cars that travel off peak may see traffic regulator delays that are largely consistent with the time intervals selected above. Those latter cars will make better surrogates.

Note that, in the multi-valued classification scenario, the quality of the surrogate may further depend on its output value. Some outputs might be more reliable whereas others more prone to misclassification. For example, a car that travels off peak but whose driver tends not to come to a full stop at stop signs may offer a good surrogate for detecting traffic lights but a bad surrogate for detecting stop signs. Hence, in the multi-valued sensing problem, the quality of a surrogate is denoted by a *vector* instead of a *scalar*. Different elements of that vector are estimated separately from data regarding the corresponding class of measurements. Other than this change, the problem of jointly deciding on values of logical quantities and quality of different surrogates remains the same. As before, the maximum likelihood estimator jointly decides on the quality of each surrogate (now represented by a vector) and the most likely values of the logical quantities estimated.

It may be useful to think of the multi-valued sensing problem as logically equivalent to multiple binary sensing problems, one per class, with the additional constraint that each multi-valued variable can ultimately have only one value (i.e., one class). Hence, if one of the constituent binary sensing problems evaluates to true, the others must evaluate to false.

Finally, note that, in crowdsourcing applications, the observations from sources don't come all at once. Instead, updates are reported over time, lending themselves better to the abstraction of a *data stream* arriving from the community of sources. A recursive EM algorithm can be used to update estimation results on the fly in view of newly arriving data.

IV. Performance

To appreciate the power of surrogate sensing, we report recent evaluation results on finding stop signs in the city of Urbana-Champaign [12]. To detect a stop sign using a phone inside a vehicle, the phone detects a 2-10 second stop as the surrogate. Stop sign claims are reported by each phone to a central data collection point.

In the experiment, 34 people (sources) were invited to participate and 1,048,572 GPS readings (around 300 hours of driving) were collected. A total of 3303 claims were generated by the phones for stop signs covering 190 different locations. The schemes compared were the basic expectation maximization (EM), as well as EM with source/measurement co-location constraints (OtO EM)[1], EM with measurement correlation constraints (DV EM)[2], and EM with both types of constraints (OTO+DV EM). Table I shows the estimation error in surrogate quality (percentage of time the surrogate is correct).

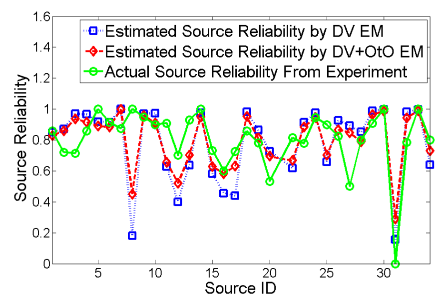

Fig. 2. Source Reliability Estimation of DV and DV+OtO EM in the Case of Stop Signs

Figure 2 shows the average source reliability estimation error, when these constraints are used. We observe that both DV EM and DV+OtO EM scheme track the source reliability very well (the estimation error of the two EM schemes improved 9.4% and 13.4% respectively compared to the regular EM scheme).

The ROC curves illustrating the trade-offs of true positives and false positives for the different EM variants are shown

[1]OtO stands for "Opportunity to Observe", since the constraint states that sources can credibly report only what they had the opportunity to observe.

[2]DV stands for "Dependent Variables".

	Regular EM	OtO EM	DV EM	DV+OtO EM
Average Surrogate Quality Estimation Error	25.34%	16.75%	15.99%	11.98%
Number of Correctly Identified Stop Signs	127	139	141	146
Number of Mis-Identified Stop Signs	25	24	29	25

TABLE I. A SURROGATE SENSING COMPARISON

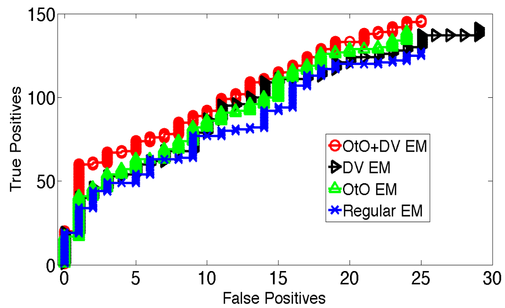

Fig. 3. ROC Curves of OtO, DV, OtO+DV EM vs Regular EM in the Case of Stop Signs

in Figure 3. Observe that the DV EM scheme finds 14 more correct stop sign locations than the Regular EM scheme. The DV+OtO EM scheme performs the best. It finds the most stop sign locations (i.e., 19 more than regular EM, 5 more than DV EM) while keeping the false positives the least (i.e., the same as regular EM and 4 less than DV EM). The detailed comparison results are given in Table I.

V. DISCUSSION

This article presented a maximum likelihood estimation framework for surrogate sensing that, among other things, exploits the physical constraints (i.e., source locations and observed variable dependencies) to improve sensing reliability. Some limitations exist that offer directions for future work.

For example, many variables change over time. The aforementioned work does not explicitly model the time dimension of the problem. “Byzantine” sources are not considered (e.g., phones that cheat in reporting their GPS coordinates). It is interesting to investigate the robustness of the scheme with respect to the percentage of cheating sources in the system. The work assumes that the joint probability distribution of dependent variables is known or can be estimated from prior knowledge. This might not be possible in some applications. Interesting questions arise regarding the convergence of the scheme. It remains to understand the precise conditions under which convergence to an optimal solution occurs.

Notwithstanding the above limitations, this initial work on surrogate sensing demonstrates the feasibility of sensing high-level quantities with complex semantics using weak surrogate sensing modalities available on the phone.

VI. CONCLUSION

This article presented a framework for exploiting quantity to promote simplicity of the sensing front-end. It suggests that the many applications can get away with cheap widely available sensors, by exploiting a smart back-end that correctly determines which data to trust. The essential problem at the back-end lies in assessing the probability of correctness of claims made by weak surrogates, which in turn requires assessment of quality of the surrogates themselves, all while exploiting physical constraints and data provenance relations. Several variants of an expectation maximization scheme were described that arrive at a maximum likelihood solution to the problem. The maximum likelihood estimator acts as a “force multiplier” for the smart-phone, allowing its common sensors to be used effectively for a much broader spectrum of mobile sensing applications than what the available sensing modalities might initially suggest.

ACKNOWLEDGEMENTS

Research reported in this paper was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement W911NF-09-2-0053, DTRA grant HDTRA1-10-1-0120, and NSF grants CNS 09-05014 and CNS 10-35736. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

REFERENCES

[1] R. Ganti, F. Ye, and H. Lei, “Mobile crowdsensing: current state and future challenges,” *Communications Magazine, IEEE*, vol. 49, pp. 32–39, November 2011.

[2] J. Aslam, S. Lim, X. Pan, and D. Rus, “City-scale traffic estimation from a roving sensor network,” in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, SenSys ’12, (New York, NY, USA), pp. 141–154, ACM, 2012.

[3] J. Biagioni, T. Gerlich, T. Merrifield, and J. Eriksson, “Easytracker: Automatic transit tracking, mapping, and arrival time prediction using smartphones,” in *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems*, SenSys ’11, (New York, NY, USA), pp. 68–81, ACM, 2011.

[4] S. Hemminki, P. Nurmi, and S. Tarkoma, “Accelerometer-based transportation mode detection on smartphones,” in *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, SenSys ’13, (New York, NY, USA), pp. 13:1–13:14, ACM, 2013.

[5] M.-R. Ra, B. Liu, T. F. La Porta, and R. Govindan, “Medusa: A programming framework for crowd-sensing applications,” in *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*, MobiSys ’12, (New York, NY, USA), pp. 337–350, ACM, 2012.

[6] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava, "Participatory sensing," in *In: Workshop on World-Sensor-Web (WSW06): Mobile Device Centric Sensor Networks and Applications*, pp. 117–134, 2006.

[7] D. Wang, L. Kaplan, H. Le, and T. Abdelzaher, "On truth discovery in social sensing: A maximum likelihood estimation approach," in *The 11th ACM/IEEE Conference on Information Processing in Sensor Networks (IPSN 12)*, April 2012.

[8] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," in *7th international conference on World Wide Web (WWW'07)*, pp. 107–117, 1998.

[9] M. Gupta and J. Han, "Heterogeneous network-based trust analysis: A survey," *SIGKDD Explor. Newsl.*, vol. 13, pp. 54–71, Aug. 2011.

[10] R. V. Hogg and A. T. Craig, *Introduction to mathematical statistics*. Prentice Hall, 1995.

[11] D. Wang, L. M. Kaplan, T. F. Abdelzaher, and C. C. Aggarwal, "On credibility estimation tradeoffs in assured social sensing," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 6, pp. 1026–1037, 2013.

[12] D. Wang, T. Abdelzaher, L. Kaplan, and R. Ganti, "Exploitation of physical constraints for reliable social sensing," in *IEEE Real-time Systems Sumposium (RTSS)*, (Vancouver, Canada), 2013.

VII. Author Biographies

Dong Wang received his Ph.D. in Computer Science from University of Illinois at Urbana Champaign in 2012. He is currently an assistant professor in the Department of Computer Science and Engineering at University of Notre Dame. Dr. Wangs research interests lie in the area of reliable social sensing, cyber-physical computing, real-time and embedded systems, and crowdsourcing applications. He received the Wing Kai Cheng Fellowship from University of Illinois in 2012 and the Best Paper Award of IEEE RTAS in 2010. He is a member of IEEE and ACM.

Tarek Abdelzaher received his Ph.D. in Computer Science from the University of Michigan in 1999. He is currently a Professor and Willett Faculty Scholar at the Department of Computer Science, the University of Illinois at Urbana Champaign. He has authored/coauthored more than 200 refereed publications in real-time computing, distributed systems, sensor networks, and control. He is an Editor-in-Chief of the Journal of Real-Time Systems, and is a recipient of the IEEE Outstanding Technical Achievement and Leadership Award in Real-time Systems (2012), the Xerox Award for Faculty Research (2011), as well as several best paper awards.

Lance M. Kaplan received his Ph.D. degree in Electrical Engineering from the University of Southern California, Los Angeles, in 1994. He is currently a researcher in the Networked Sensing and Fusion branch of the U.S. Army Research Laboratory. Dr. Kaplan is on the Board of Governors of the IEEE AES Society and on the Board of Directors of the International Society of Information Fusion. He serves as Editor-In-Chief for the IEEE Transactions on Aerospace and Electronic Systems and is a three time recipient of the Clark Atlanta University Instructional Excellence Award.