This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2784358, IEEE Transactions on Signal Processing

Experiments and Models for Decision Fusion by Humans in Inference Networks

Aditya Vempaty, *Member, IEEE*, Lav R. Varshney, *Senior Member, IEEE*, Gregory J. Koop, Amy H. Criss, and Pramod K. Varshney, *Fellow, IEEE*

***Abstract*—With the advent of the Internet of Things (IoT) and a rapid deployment of smart devices and wireless sensor networks (WSNs), humans interact extensively with machine data. These human decision makers use sensors that provide information through a sociotechnical network. The sensors can be other human users or they can be IoT devices. The decision makers themselves are also part of the network, and there is a need to understand how they will behave. In this paper, the decision fusion behavior of humans is analyzed on the basis of behavioral experiments. The data collected from these experiments demonstrate that people perform decision fusion in a stochastic manner dependent on various factors, unlike machines that perform this task in a deterministic manner. A Bayesian hierarchical model is developed to characterize the observed stochastic human behavior. This hierarchical model captures the differences observed in people at individual, crowd, and population levels. The implications of such a model on designing large-scale inference systems are presented by developing optimal decision fusion trees with both human and machine agents.**

***Index Terms*—human behavior modeling, decision fusion, Bayesian hierarchical modeling, sociotechnical networks**

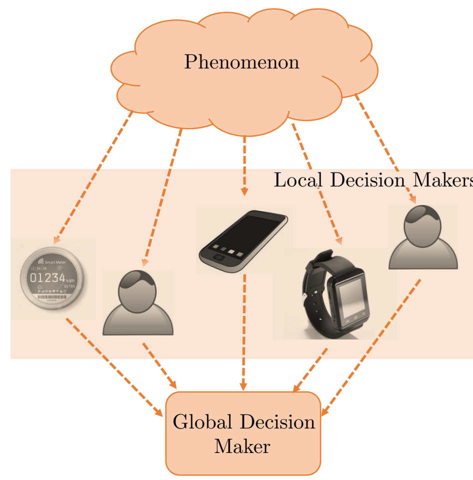

Fig. 1. System model consisting of local decision makers and a global decision maker.

I. Introduction

Sociotechnical networks capture the interaction of human behavior with society's complex infrastructures. The optimal design of such networks considers human, social, and organizational factors, besides technical ones [2]. The information flow within such systems is supported by the technical part such as a sensor network. The presence of humans in the system, who can take actions, affects both the sociotechnical and the technical parts of the system [3]–[7]. For example, human decisions determine movement patterns for many mobile devices, which in turn impact load and connectivity. These same decisions impact the ability of people to observe a given phenomenon. While human actions are not completely deterministic, they can be predictable. Consider a crowdsensing system where humans make decisions based on local decisions from information sources such as other people or IoT devices. These decisions are then sent over (imperfect) channels to a fusion center for decision making. For such systems, it is important to develop efficient techniques to model human behavior while fusing decisions. To characterize how people fuse multiple decisions to make their own decisions, this work presents behavioral experiments for this task and develops a Bayesian hierarchical model that describes this behavior. Further, making use of our Bayesian hierarchical model of human behavior, we develop optimal decision fusion trees with both humans and IoT devices. In particular, we incorporate the randomness associated with human behavior into the design of fusion rules and show the improvement in performance by using such rules.

Decision fusion is the process of integrating decisions made by multiple entities about the same phenomenon into a single final decision. The typical framework of parallel decision fusion is shown in Fig. 1, where a set of local decision makers (LDMs) observe a phenomenon and make decisions regarding its presence or absence (Yes/No binary decisions). These local decisions are received by a global decision maker (GDM) who fuses the received data to make the final decision.

This work was supported in part by NSF under Grants CCF-1623821 and CCF-1717530. The material in this paper was presented in part at the 3rd IEEE Global Conference on Signal and Information Processing (GlobalSIP), December 2015 [1].

A. Vempaty was with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA. He is now with the IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598 USA (e-mail: avempat@us.ibm.com).

L. R. Varshney is with the Department of Electrical and Computer Engineering and with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: varshney@illinois.edu).

G. J. Koop is with the Department of Psychology, Eastern Mennonite University (e-mail: gregory.koop@emu.edu).

A. H. Criss is with the Department of Psychology, Syracuse University, Syracuse, NY 13244 USA (e-mail: acriss@syr.edu).

P. K. Varshney is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail: varshney@syr.edu).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2784358, IEEE Transactions on Signal Processing

In the signal processing literature, such problems have been extensively studied when all the decision makers are machines [8]–[11] and optimal decision rules for both local decision makers and global decision maker have been designed under various assumptions [8], [12], [13]. When the global decision maker is using an optimized fusion rule but the local decision makers are humans, the above framework addresses the paradigm of crowdsensing for distributed inference tasks [14]–[17]. In such systems, one can analyze the system performance and design simple easy-to-perform tasks to improve the overall performance of the system [18]. To engineer networks where the *global decision maker is also a human* which arises in the sociotechnical systems described above, it is of interest to understand how people fuse decisions. In this paper, based on experimental results, we develop a particular bounded rationality model (cf. [19]).

Understanding the human-decision making process using signal processing techniques and its effect on sociotechnical systems has gained increasing interest among researchers [20]–[25], especially due to the advent of social networks. In [21], Rhim et al. study collaborative distributed hypothesis testing by a group of agents who have knowledge of quantized prior probabilities [20], drawn from an ensemble. They study the effect of such quantization of prior probabilities on distributed detection performance. Wimalajeewa and Varshney also consider the problem of collaborative human decision making but model the humans as decision makers who follow threshold-based decision schemes and model the thresholds as random variables [22]. The performance of such systems is characterized in terms of probability of error and the optimal statistical parameters of the threshold distributions are analytically derived. In contrast, [25] considers the framework where the human agents make sequential decisions where the next agent's decision depends on their private observation and the previous agent's decision. The performance of such a social learning framework is contrasted to the typical distributed decision making framework under different scenarios. While our paper deals with the same framework when decision making agents are all humans, the specific focus here is on the case where a human global decision maker is fusing decisions from multiple human local decision makers.

In this paper, a similar signal processing methodology is applied to understand the process of decision fusion by humans. The problem of fusing multiple human decisions has been investigated in different contexts in the psychology literature (see [26], [27], and references therein). Such a framework is also very similar to problems in social choice theory and voting. These systems have been studied under idealizations of human behavior, including likelihood ratio tests with Bayes-optimal thresholds and deterministic, optimized, symmetric decision fusion. However, past literature and our new experimental data show that human behavior is not generally deterministic and so people do not perform Bayes-optimal decision fusion. We find through experimentation that none of the five reasonable fusion rules considered here provide a good match to human behavior. Therefore, we propose a Bayesian hierarchical model [28] to replicate the behavior of a population of human decision aggregators. The model is a symmetric perturbation of one of three fusion rules (to be detailed later). Note that the model does not necessarily capture human behavior at the level of individual choices, but instead replicates the randomness associated with human decision making through a generative model. This helps in the design of large-scale systems that are affected by such human behavior. In such cases, it is useful to know how a population performs fusion because there may be further downstream decision making that can be optimized based on an understanding of how the intermediate decisions have been made. We demonstrate the potential improvement quantitatively using analytical expressions and simulations.

This paper builds on the preliminary work reported in [1]. Significantly more experimental data (almost three times the preliminary work) was collected for this paper, resulting in more accurate results. Besides the experimental data, this paper also improves the Bayesian hierarchical model used in [1] to accommodate multiple fusion rules by humans. Our preliminary work focused only on the optimal fusion rule (Chair-Varshney rule). However, further discussions within the research team, which includes psychologists, revealed several other sub-optimal fusion rules that are used by humans (see Sec. III-A). Therefore, the model was enhanced by adding another dimension of stochasticity to accommodate the existence of different fusion rules. The larger experimental data, the enhanced models, and more accurate results make this work a complete version of the preliminary work reported in [1].

The remainder of the paper is organized as follows. In Sec. II, we describe psychology experiments designed to understand human decision fusion. Preliminary analysis of the collected data is performed in Sec. III by comparing the observed decisions with several popularly used fusion rules. After establishing that existing decision fusion models cannot explain the human behavior, in Sec. IV we build a Bayesian hierarchical model to explain the observed behavior. In Sec. V, we discuss its implications by demonstrating its effect on the design of large-scale hierarchical sociotechnical systems, consisting of multiple human decision fusion components. We conclude the paper in Sec. VI.

II. Experiments

To understand decision fusion behavior in humans, experiments replicating the process of Fig. 1 were designed. Human subjects consisting of undergraduate students at Syracuse University were enrolled for this task.[1] The experiment consisted of data collection in two stages: the first stage models local decision making and the second stage models data fusion. The experiment is that of a memory-based task and is described as follows.

A. Stage 1: Local decision making

a) Participants: A total of 45 introductory psychology students from Syracuse University performed the first stage of the experiment that models local decision making. All participants received partial fulfillment of course requirements for their participation.

[1]The necessary IRB approval was obtained before conducting the experiments.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2784358, IEEE Transactions on Signal Processing

b) Stimulus materials: A study list $\mathcal{D}$ containing 100 English words ranging in length from 5 to 11 letters (median = 7), and ranging between 8.41 and 12.17 log frequency (mean = 10.33, standard deviation = 0.93) in the Hyperspace Analog to Language Corpus (HAL) [29] was provided to the participants. A test list $\mathcal{S}$ containing words in $\mathcal{D}$, and an additional 100 distractor words $\mathcal{N}$ was prepared ($\mathcal{S} = \mathcal{D} \cup \mathcal{N}$). These distractor words in $\mathcal{N}$ were between 5 and 12 letters in length (median = 7), and ranged from 8.10 to 13.27 log frequency (mean = 10.34, standard deviation = 0.94) in the HAL Corpus.

c) Procedure: After providing informed consent, participants were seated in individual testing booths and instructed that they would study a series of words and then have their memory for those words tested. During the study phase, participants were asked to indicate whether each of the presented words $s \in \mathcal{S}$ belonged to the previously memorized target set ($s \in \mathcal{D}$) or the unseen distractor set ($s \in \mathcal{N}$). Participants were required to make this judgment within 6 seconds, or else the trial was discarded and the next item would appear automatically. Each participant completed 200 such trials. The order in which these 200 trials unfolded was randomized for each participant.

B. Stage 2: Global decision making

a) Participants: A total of 60 introductory psychology students from Syracuse University participated in the experiment. All participants received partial fulfillment of course requirements for their participation.

b) Stimulus materials: The stimuli used in the second stage are the recognition judgments provided by participants in Stage 1. For example, in the first stage, participants had their memory tested for the word *Project.* A trial in the second stage showed the word *Project* as well as the recognition judgments from a varying number of "sources" (i.e., participants in the first stage). For each of these sources, participants saw three pieces of information: sources' decisions, accuracies, and bias values. See Fig. 2 for an example. Source accuracy is defined as the proportion of correct answers (i.e., "hits" on target trials and "correct rejections" on distractor trials) over the course of the experiment (excluding trials on which no answer was given, as described above). The bias values (the far right column of Fig. 2) represented how frequently a source gave a "yes" response across both target and distractor trials.

c) Procedure: Upon arrival to the test setting, participants provided informed consent and received instructions about the task. All the participants were told that earlier in the semester, participants like themselves had completed a recognition memory task (i.e., Stage 1). We simply asked participants to try and identify whether or not a word was truly studied on the basis of responses from participants in the previous experiment. The decision task consisted of 200 trials, where participants were provided with information from a varying number of sources (N) from Stage 1 (these were real participants from Stage 1). Participants saw 2, 5, 10, or 20 source judgments. The number of sources presented on each trial was randomized over the course of the experiment, with the lone constraint that there were 50 trials of each type. On each trial, participants were asked to provide a *yes* or *no* response to the question "Do you think this item was actually studied?" Responses were indicated by a single key press ("z" or "/?"). Response keys were counterbalanced across participants. Finally, after making their judgment, participants were also asked to assess their confidence in that judgment (1 = low; 3 = high). After providing this confidence judgment, participants proceeded to the next trial. Each datapoint of the resulting dataset has the following information: word s, true hypothesis of s ($s \in \mathcal{D}$ or $s \in \mathcal{N}$), number of sources for this particular task (N), sources' decisions ($u_1, \ldots, u_N$), sources' accuracy ($a_1, \ldots, a_N$) and bias values ($b_1, \ldots, b_N$), and the fused decision reported by the global decision maker (d).

Project

	Decision	ACC	Pr("yes")
Source A	Yes	0.66	0.57
Source B	Yes	0.80	0.36
Source C	Yes	0.73	0.68
Source D	No	0.87	0.54
Source E	No	0.80	0.40

z **YES** Do you think this item was actually studied? /? **NO**

Fig. 2. Example trial where participants in Stage 2 had to decide if the word *Project* is a part of the study list $\mathcal{S}$ or the distractor list $\mathcal{N}$ using information provided by participants from Stage 1.

Note that for an accurate understanding of human decision behavior, it was necessary to actually perform the first stage of the experiment with human subjects instead of randomly generating local decisions for the humans in the second stage. This is because humans have item-specific bias and accuracy levels [30], [31]. For example, people tend to believe that they remember negative arousing words (such as *murder*) regardless of whether or not they actually studied the word and people tend to have high accuracy (high hit rates and low false alarm rates) for uncommon words (such as *ire*). If decisions were randomly generated for the second stage, then the data would not reflect these factors and would probably be distrusted (consciously or unconsciously) by the decision makers. In a typical experiment, these problems are avoided by fully randomizing assignment of words to $\mathcal{D}$ or $\mathcal{N}$ status. For this experiment, we did not randomize so that we could have N decisions under (approximately) identical circumstances (same $\mathcal{D}$ and $\mathcal{N}$) with the added randomness of individual differences among people. However, as noted before, the order of presentation of trials was randomized in the second stage since that affects performance [32], [33].

III. Preliminary data analysis

This section presents a preliminary analysis of the collected data.[2] Decisions made by humans in the experiments are compared against some known fusion rules. First, traditional decision fusion rules are presented in Sec. III-A. The decisions of these traditional decision rules are then compared to the observed decisions of the humans in Sec. III-B.

A. Fusion rules

1) Optimal fusion rule (CV): When the sources' reliabilities are known, optimal decision fusion is achieved by the Chair-Varshney (CV) rule [12]. Represent the "Yes/No" decisions of the ith local decision maker as

$$u_i = \begin{cases} +1, & \text{if the decision is "Yes"}, \\ -1, & \text{if the decision is "No"}. \end{cases} \tag{1}$$

After receiving the N decisions $\mathbf{u} = [u_1, \ldots, u_N]$, the global decision $u_0 \in \{-1, +1\}$ is made as follows:

$$d = \begin{cases} +1, & \text{if } m_0 + \sum_{i=1}^{N} m_i u_i > 0, \\ -1, & \text{otherwise}, \end{cases} \tag{2}$$

where $m_0 = \log \frac{P_1}{1-P_1}$,

$$m_i = \begin{cases} \log \frac{1-P_{M,i}}{P_{F,i}}, & \text{if } u_i = +1, \\ \log \frac{1-P_{F,i}}{P_{M,i}}, & \text{if } u_i = -1, \end{cases} \tag{3}$$

for $i = 1, \ldots, N$ is defined as the *reliability* of a decision, and P_1 is the prior probability that the underlying hypothesis is "Yes" (+1), $P_{M,i}$, $P_{F,i}$ represent the probability of missed detection and false alarm respectively, of the ith decision maker.

2) Most accurate decision (MAD): The *most accurate* decision rule is a heuristic decision rule that has been described in human decision making literature. It is defined as follows: $d = u_a$ where

$$a = \arg\max_i a_i, \tag{4}$$

and a_i is the accuracy of the ith local decision maker. In terms of missed detection and false alarm probabilities, this is given as

$$a_i = P_0(1 - P_{F,i}) + P_1(1 - P_{M,i}). \tag{5}$$

This decision rule only depends on the accuracy values of the local decision makers and is therefore believed to be a strong heuristic used by humans especially when the number of decisions presented for fusion (N) is large.

3) Most reliable decision (MRD): The *most reliable* decision rule is another heuristic decision rule considered in this paper. It is defined as $d = u_\rho$ where

$$\rho = \arg\max_i m_i,$$

and m_i is the reliability of ith local decision maker given by (3). This decision rule depends on both accuracy and bias values of the local decision makers.

[2]This data is available on the Open Science Framework at https://osf.io/a7pgz/.

TABLE I
Mean $\pm$ standard deviation of match values for different values of N and for different rules

N	CV	MAD	MRD	CCV-0.9	MAJ
2	0.80±0.17	0.81±0.18	0.80±0.17	0.57±0.12	0.47±0.09
5	0.83±0.18	0.76±0.14	0.75±0.14	0.75±0.18	0.46±0.10
10	0.83±0.18	0.75±0.14	0.74±0.14	0.79±0.17	0.46±0.09
20	0.83±0.18	0.73±0.13	0.73±0.13	0.82±0.17	0.45±0.09

4) Censored CV rule (CCV-τ): The *censored CV* decision rule with parameter τ is a censored version of the CV rule of Sec. III-A1 that may be used by humans when the number of sources is large. It is mathematically given as

$$d = \begin{cases} +1, & \text{if } m_0 + \sum_{i=1}^{N} \tilde{m}_i u_i > 0, \\ -1, & \text{otherwise}, \end{cases} \tag{6}$$

where u_i is given by (1) and

$$\tilde{m}_i = \begin{cases} m_i, & \text{if } m_i \geq \tau, \\ 0, & \text{otherwise}. \end{cases} \tag{7}$$

Here, τ is the censoring threshold that determines when a particular decision is reliable and therefore, should be considered in the decision making process.

5) Majority rule (MAJ): The *majority* rule is a very common decision rule used in practice, especially when the accuracy or bias values of the local decision makers are unavailable. It is given as

$$d = \begin{cases} +1, & \text{if } \sum_{i=1}^{N} u_i > 0, \\ -1, & \text{otherwise}, \end{cases} \tag{8}$$

where u_i is given by (1).

B. Comparison of fusion rules

Before building a model of how humans fuse data, we compare the experimental data with the fusion rules described in Sec. III-A. For this purpose, final decisions of $T = 60$ human global decision makers are compared with the decision from the fusion rules described in Sec. III-A. Note that in our setup, $P_1 = 0.5$, implying $m_0 = 0$. Each human subject at the second stage typically performed 100 trials, 25 each with $N = 2, 5, 10, 20$. The final decisions made by the humans are compared with the decisions made by the fusion rules with the same input. The fraction of times that a decision maker i's decision matches the decision of fusion rule r with the same input data is defined as the *match value* $p_{i,r}$ of the ith decision maker with rth rule. Table I shows the mean of match values across all human subjects for each of the fusion rules with varying number of local decision makers.

As we can observe from Table I, the average match value improves with increasing number of sources for the CV-based (CV and CCV-τ) rules but not necessarily for the other rules. Also, on comparing the individual match values $p_{i,r}$ across the rules, we observed that the CCV and the MAJ rules were never the best for any of the 60 individuals for any value of N. However, the other three rules were better for some individuals. For example, when $N = 2$, the CV rule was best among all rules for participant id 61, whereas the MAD rule was the best one for most of the individuals. On the other hand, for the same participant with id 61, when $N = 5$, the CV rule had the highest match value. Also, the variability of match values is very high, with some participants having a match value close to 1, while some having as low as 0.32. Therefore, a single decision fusion rule cannot capture every human's behavior at every time instant. In the following, we develop a Bayesian hierarchical model to represent the observed human behavior.

IV. Bayesian hierarchical model

In this section, a Bayesian hierarchical model is developed which characterizes the human behavior when fusing multiple decisions. This model encapsulates the variability among human behavior observed at an individual level, crowd level, and population level.

A. Description of model

From the preliminary data analysis of the previous section, we observed that no single rule perfectly characterizes the behavior for all individuals. Consider a discrete set of fusion rules $\mathcal{R}$. Then, one can model an individual to be using a fixed fusion rule $r_i = j \in \mathcal{R}$ and a fixed match value p_i. On the other hand, this rule r_i and the match value p_i differs for every individual. Even among all individuals who use the same fusion rule j, the match value differs. This behavior can be captured by modeling the fusion rule r_i as a random variable following a distribution $f_r(\cdot)$ with support set $\mathcal{R}$ and the match value p_i as a random variable with distribution $f_{p,r}(\cdot)$. Such a model captures the individual differences in humans while fusing multiple decisions. As mentioned before, the differences among humans can be at multiple levels: individual level, crowd level, and population level. The individual-level decision model is described below (Fig. 3):

- A deterministic decision v is determined using the fusion rule j, which is fixed for an individual.
- The individual's final decision is determined by flipping the deterministic decision v with probability $(1-p)$ where p is the individual's match value.[3]

Therefore, the final decision is now given by:

$$d = \begin{cases} v, & \text{with probability } p, \\ 1-v, & \text{with probability } 1-p. \end{cases} \tag{9}$$

This randomness in human decision making can be attributed to the fact that human perception and encoding (of the stimulus) is subject to uncertainty. Therefore, rather than implementing a mechanistic account of that, we characterize the randomness by introducing noise in the decision for simplicity.

Moving another step higher in the hierarchy, at the crowd level, every individual has their fixed fusion rule $r_i = j$ that is determined by sampling from distribution $f_r(\cdot)$ and the match value p_i for the individual is sampled from a distribution $f_{p,r}(\cdot)$. These distributions $f_r(\cdot)$ and $f_{p,r}(\cdot)$ are determined by fitting a model to experimental data of Sec. II. For our models, we consider $f_r(\cdot)$ to be a categorical distribution with parameters $\mathbf{q}$ where q_j denotes the probability of choosing fusion rule j and $\sum_j q_j = 1$. The distribution $f_{p,r}(\cdot)$ is modeled to be a beta distribution with parameters α_j and β_j which depend on the fusion rule j. Let $\alpha = [\alpha_1, \ldots, \alpha_j, \ldots, \alpha_R]$ and $\beta = [\beta_1, \ldots, \beta_j, \ldots, \beta_R]$ where $R = |\mathcal{R}|$ is the total number of fusion rules. The parameters $\mathbf{q}$, α, and β correspond to the crowd parameters that serve as hyperparameters for r and p.

As we shall see later, the values of the hyperparameters $\mathbf{q}$, α, and β themselves depend on the crowd considered, i.e., they depend on the number of sources, whether they are college students or online participants, the demographics of the participants, etc. This takes us to the higher level in the model where these values of $\mathbf{q}$, α, and β, or in other words, the distributions $f_r(\cdot)$ and $f_{p,r}(\cdot)$ themselves depend on the underlying crowd chosen for the task. Different crowds would have different values of $\mathbf{q}$, α, and β. Hidden variables like demographics, motivation, etc. can affect the parameters of the randomized decision rule model discussed above. Therefore, continuing on the Bayesian modeling approach, these parameters $\mathbf{q}$, α, and β can be modeled as random variables sampled from a distribution with parameters $\mathcal{P}$ (population parameters). The distribution of $\mathbf{q}$ could be the conjugate prior of categorical distribution, i.e., the Dirichlet distribution. Similarly, the distribution α and β can be the conjugate prior of the beta distribution, which exists since the beta distribution falls under the family of exponential distributions. In this case, the parameters of the Dirichlet distribution and the parameters of the conjugate prior of Beta distribution serve as the population parameters. Population parameters govern the entire population as a whole from which different sets of crowds are sampled. This complete model can be captured by Fig. 4.

B. Model inference

In this section, we infer the parameters of the model using data collected in Sec. II. From our observations, we saw that the CCV rule and the MAJ rule were not the best fit rules for any individuals in our dataset. Therefore, we consider $R = 3$ and consider the rules to be [CV, MAD, MRD]. The optimal approach of using a joint maximum likelihood approach would require the knowledge of the latent variable, i.e., the knowledge of fusion rule being used by each individual. We first infer $\mathbf{q}$ followed by the parameters (α_j, β_j) as follows. Note that these parameters can also be jointly estimated using an EM-based method.

1) Inferring $\mathbf{q}$*:* The rule selection parameter $\mathbf{q}$ is inferred using a maximum likelihood estimate as follows. We first determine the match values corresponding to every rule for every individual. Represent the match value of individual i with rule j as $p_{i,j}$. Let $\tilde{p}_i$ represent the maximum value among all $p_{i,j}$ for a fixed i, i.e.,

$$\tilde{p}_i = \max_j p_{i,j}. \tag{10}$$

[3] A match value of $p > 0$ in our model captures the model for limited rationality.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2784358, IEEE Transactions on Signal Processing

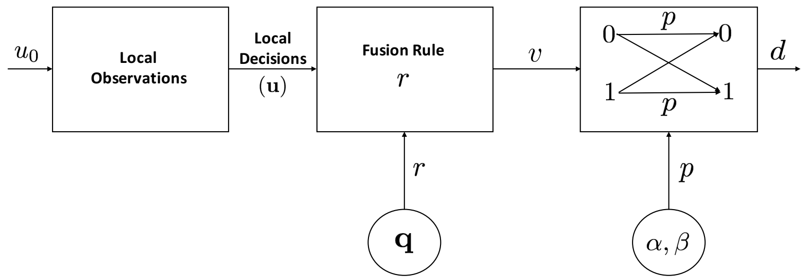

Fig. 3. Proposed 2-step model where the first step determines a deterministic decision using rule r and the second step models the randomness of an individual human's decisions by using a match value p. Here $\mathbf{q}$, α and β are hyperparameters that capture the randomness of r and p, among multiple individuals at the crowd level.

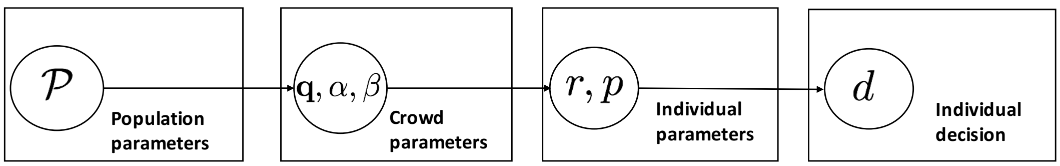

Fig. 4. Bayesian hierarchical model of decision fusion by humans using the plate notation of representing variables in a graphical model.

Now, let $0 \leq T_j \leq T$ represent the number of individuals among the T individuals for whom $\tilde{p}_i = p_{i,j}$. This is the empirical number of individuals that follow rule j. The estimate of $\mathbf{q}$ is then determined as a (normalized) version of

$$\hat{q}_j = \frac{T_j}{T}. \tag{11}$$

An additional normalization step might be needed since multiple rules can result in the same match value that is equal to the maximum one. Normalization ensures that the constraint $\sum_j q_j = 1$ is satisfied.

2) Inferring α and β: The parameters of the beta distribution are identified as follows. For learning α_j and β_j, we only consider the T_j individuals who follow rule j. Due to the limited number of data points, a bootstrap model is used for data fitting, where $t = 0.7T_j$ data points among the total T_j data points are randomly selected for which a beta distribution is fit. This process is repeated $N_{mc} = 1000$ times. If α_k and β_k represent the parameters from the kth trial, the final parameters are decided by taking an average of these parameters.

3) Inference results: The results are compiled in Table II and Figs. 5-8. Table II presents all the inferred parameter values for different values of N (the number of sources). As we can observe, more individuals followed sub-optimal fusion rules for lower values of N and the optimal CV rule for higher number of sources. Also, the mean of the match value, $E[p] = \alpha/(\alpha+\beta)$ increases with an increase in N. To gain further insights on how the distribution of the match value varies for different rules and for different values of N, we plot the distributions in Figs. 5-8. An interesting observation is that the distribution $f_{p,CVrule}(\cdot)$ has increasing mean and shifts to the right with increase in N, while the distributions of other rules (MAD and MRD) do not necessarily follow such a trend. Also, the distribution $f_{p,CVrule}(\cdot)$ corresponding to the optimal CV rule has constant shape and robust parameters with increasing N while the distributions for the MAD and MRD

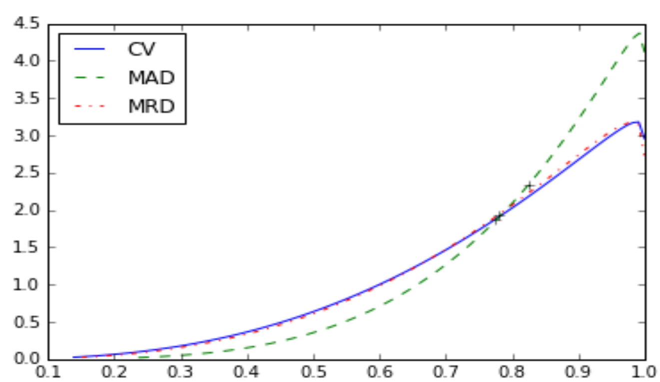

Fig. 5. Distribution $f_{p,j}(\cdot)$ of match value p for different fusion rules j when $N = 2$, based on data fitting. The mean value is also highlighted. Cross represents mean value of the distribution.

rules are less robust to the number of sources. This could be an artifact of limited data as there were relatively lesser data points for these rules in comparison to CV rule. This intuition will be further explored in the future by collecting higher number of data points.

TABLE II
PARAMETERS FOR DIFFERENT VALUES OF N. THE PARAMETERS FOR DIFFERENT RULES ARE PRESENTED IN THE ORDER [CV, MAD, MRD].

N	$\mathbf{q}$	α	β
2	[0.26, 0.47, 0.27]	[3.52, 4.82, 3.70]	[1.03, 1.03, 1.05]
5	[0.84, 0.10, 0.06]	[5.35, 53.41, 77.69]	[1.06, 13.11, 53.86]
10	[0.81, 0.10, 0.09]	[5.45, 28.24, 45.92]	[1.01, 15.68, 23.25]
20	[0.80 , 0.09, 0.11]	[6.04, 18.28, 29.71]	[0.98, 8.55, 15.76]

From the proposed model, it is clear that for a complete study, one has to repeat human subject experiments with different crowds, to determine the population parameters and

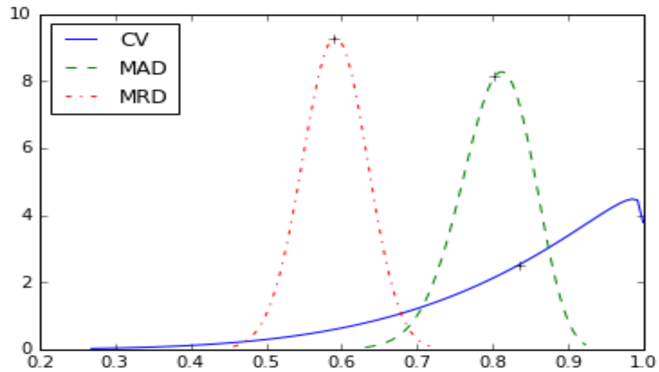

Fig. 6. Distribution $f_{p,j}(\cdot)$ of match value p for different fusion rules j when $N = 5$, based on data fitting. The mean value is also highlighted. Cross represents mean value of the distribution.

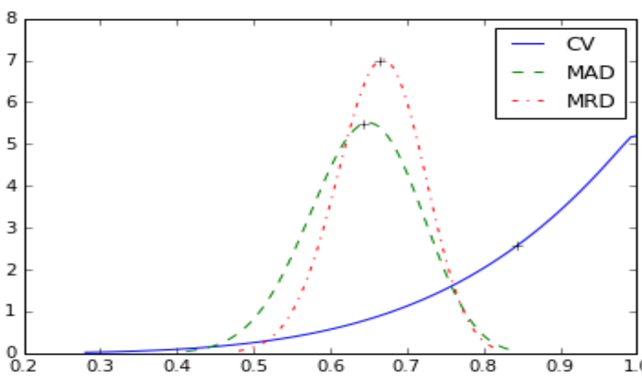

Fig. 7. Distribution $f_{p,j}(\cdot)$ of match value p for different fusion rules j when $N = 10$, based on data fitting. The mean value is also highlighted. Cross represents mean value of the distribution.

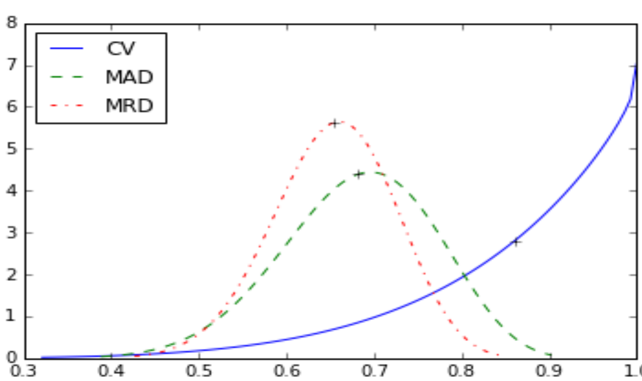

Fig. 8. Distribution $f_{p,j}(\cdot)$ of match value p for different fusion rules j when $N = 20$, based on data fitting. The mean value is also highlighted. Cross represents mean value of the distribution.

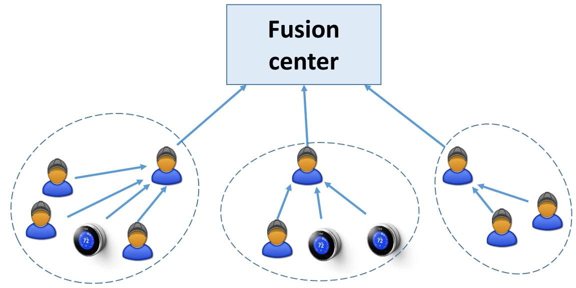

Fig. 9. Hierarchical system consisting of human decision fusion components.

their effect on the crowd parameters $\mathbf{q}$, α, and β. For example, one might get different results from online participants, such as crowd workers as compared to a group of college students [34]. Also, it has been found that age of the crowd (older vs. younger adults), or disease conditions of typical vs. atypical crowds (PTSD, dementia, Alzheimer's, etc.), might give different results [35]. From the experiments, an ensemble of parameters can be determined, which will help us in getting population-level insight into individual differences regarding how people fuse decisions. Such a hierarchical model can be used for understanding and designing larger signal processing systems that have a human decision fusion component such as distributed detection systems [8], [36] where each agent is not a single cognitive agent, but rather a human-based decision fusion system (Fig. 9). Also, cognitive agents in such systems may be drawn from a specialized sub-population.

V. Optimal Design of Sociotechnical Networks

As described in Sec. I, crowdsensing with human decision fusion components plays a key role in sociotechnical systems. Here we consider designing such sociotechnical systems with machines and with humans, as modeled through our Bayesian hierarchical framework.[4] Consider a system like Fig. 9 where multiple levels of decision makers are present in the system with human decision makers fusing data from multiple subordinate agents (humans or machines) before sending their fused observations to a final fusion center via imperfect channels. If these last-level agents were IoT devices rather than humans, one could use the optimal fusion rule to fuse the data [12]. Note that this optimal fusion rule weighs the decisions with their reliabilities which are deterministically known. However, when the final fusion center receives data from humans and via imperfect mobile channels, one needs to use the Bayesian hierarchical model of human decision fusers along with the channel effects to design the fusion rule at the fusion center.[5]

Considering the Bayesian formulation, the optimal fusion rule at the fusion center is developed by adopting a methodology similar to [12]. Let the phenomenon of interest be a binary hypothesis testing problem with prior probabilities $P(H_0) = P_0$ and $P(H_1) = P_1 = 1 - P_0$. Assume that the fusion center receives decisions from M human decision

[4]Note that these intermediate agents implicitly have the goal of being right in contrast to the goal of being informative to later-acting agents [37].

[5]Note there are two kinds of hierarchies considered herein: the Bayesian hierarchy for human modeling and tree hierarchy of decision making.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2784358, IEEE Transactions on Signal Processing

fusion components. We represent the received decisions by $r_i \in \{-1,+1\}$ and the decisions made by the decision fusion component as $d_i \in \{-1,+1\}$, for $i \in \{1,\ldots,M\}$, where $r_i = \pm 1$, if the decision received from the ith component is H_1 or H_0, respectively. The fusion center makes the final decision $r_0 = f(r_1,\ldots,r_M)$ using the M decisions based on the fusion rule $f(\cdot)$. The goal is to design the optimal fusion rule $f(\cdot)$ based on the hierarchical decision making model of the components as discussed above (see Fig. 4) and the channel model between the decision fusion component and the fusion center. Consider the channels between the decision fusion component and the fusion center to be binary symmetric channels (BSC) with crossover probability p_b.

The optimal decision rule that minimizes the probability of error at the fusion center is given by the following likelihood ratio test[6]

$$\frac{P(r_1,\ldots,r_M|H_1)}{P(r_1,\ldots,r_M|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0}{P_1}, \tag{12}$$

or equivalently,

$$\log\frac{P(H_1|r_1,\ldots,r_M)}{P(H_0|r_1,\ldots,r_M)} \underset{H_0}{\overset{H_1}{\gtrless}} 0. \tag{13}$$

This optimal fusion rule can be written as

$$\log\frac{P_1}{P_0}+\sum_{\mathcal{S}_\oplus}\log\frac{P(r_i=+1|H_1)}{P(r_i=+1|H_0)}+\sum_{\mathcal{S}_\ominus}\log\frac{P(r_i=-1|H_1)}{P(r_i=-1|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} 0, \tag{14}$$

where $\mathcal{S}_\oplus$ and $\mathcal{S}_\ominus$ are the sets of all components whose received decision is $r_i = +1$ or $r_i = -1$, respectively.

The terms in (14) can be further simplified as

$$\begin{aligned}
&P(r_i=+1|H_1)\\
&= P(r_i=+1|d_i=1,H_1)P(d_i=+1|H_1)\\
&\quad + P(r_i=+1|d_i=-1,H_1)P(d_i=-1|H_1)\\
&= (1-p_b)P(d_i=+1|H_1)+p_bP(d_i=-1|H_1).
\end{aligned} \tag{15}$$

Here, $P(d_i = +1|H_1)$ is the probability that the ith decision fusion component made a decision $d_i = +1$ when the true hypothesis is H_1 and is determined using the Bayesian hierarchical model as

$$\begin{aligned}
&P(d_i=+1|H_1)\\
&= P(d_i=+1,d_{i,j}=+1|H_1)+P(d_i=+1,d_{i,j}=-1|H_1)\\
&= P(d_i=+1|d_{i,j}=+1)P(d_{i,j}=+1|H_1)\\
&\quad + P(d_i=+1|d_{i,j}=-1)P(d_{i,j}=-1|H_1)\\
&= p_iP_{d,i,j}+(1-p_i)(1-P_{d,i,j})\\
&= 1-p_i-P_{d,i,j}+2p_iP_{d,i,j}
\end{aligned} \tag{16}$$

where $d_{i,j} \in \{-1,+1\}$ is the decision that the ith human fusion center would make using its fusion rule j, p_i is the match value of the ith human corresponding to his/her rule j,[7] and $P_{d,i,j} \triangleq P(d_{i,j} = +1|H_1)$ is the probability of detection of ith decision fusion component using fusion rule j. Similarly, the expressions for $P(d_i = +1|H_0)$, $P(d_i = -1|H_1)$, and $P(d_i = -1|H_0)$ can be derived as a function of $P_{f,i,j} \triangleq P(d_{i,j} = +1|H_0)$ (false alarm probability) and are given as

$$P(d_i=+1|H_0) = 1-p_i-P_{f,i,j}+2p_iP_{f,i,j}, \tag{17}$$

$$P(d_i=-1|H_1) = p_i+P_{d,i,j}-2p_iP_{d,i,j}, \tag{18}$$

and

$$P(d_i=-1|H_0) = p_i+P_{f,i,j}-2p_iP_{f,i,j}. \tag{19}$$

Using (15)–(19), the optimal fusion rule (14) becomes

$$\begin{aligned}
&\log\frac{P_1}{P_0}+\sum_{\mathcal{S}_\oplus}\log\frac{p_b+(1-2p_b)(1-p_i-P_{d,i,j}+2p_iP_{d,i,j})}{p_b+(1-2p_b)(1-p_i-P_{f,i,j}+2p_iP_{f,i,j})}\\
&\qquad+\sum_{\mathcal{S}_\ominus}\log\frac{p_b+(1-2p_b)(p_i+P_{d,i,j}-2p_iP_{d,i,j})}{p_b+(1-2p_b)(p_i+P_{f,i,j}-2p_iP_{f,i,j})} \underset{H_0}{\overset{H_1}{\gtrless}} 0.
\end{aligned}$$

Note that the above expression requires the knowledge of every individual decision fusion component's rule j and match value p_i. When this knowledge is not available, but the crowd parameters $\mathbf{q}$, α, and β are known (refer to Fig. 4), (16) becomes

$$\begin{aligned}
P(d_i=+1|H_1) &= \sum_j q_jP(d_i=+1|j,H_1),\\
&= \sum_j q_j\int_p P(d_i=+1|j,p_i,H_1)f_{p,j}(p)dp\\
&= \sum_j q_j\Bigg(p_b+(1-2p_b)\\
&\quad\left(1-\frac{\alpha_j}{\alpha_j+\beta_j}-P_{d,i,j}+2\frac{\alpha_jP_{d,i,j}}{\alpha_j+\beta_j}\right)\Bigg)\\
&= p_b+(1-2p_b)\\
&\quad\left(1-\sum_j q_j\mu_j-\sum_j q_jP_{d,i,j}+2\sum_j q_j\mu_jP_{d,i,j}\right).
\end{aligned}$$

where $\mu_j \triangleq \frac{\alpha_j}{\alpha_j+\beta_j}$. Similarly the expressions in (17)–(19) change accordingly.

Therefore, when all the decision fusion components are identical (same number of sources, identically distributed sources, identically distribution fusion rule selection, etc.), then the optimal fusion rule becomes a K-out-of-M rule. The optimal K is easy to derive and is given by

$$K^* = \left\lceil\frac{\log\frac{P_0}{P_1}-M\log a^*_\ominus}{\log\frac{a^*_\oplus}{a^*_\ominus}}\right\rceil, \tag{20}$$

where

$$a^*_\oplus = \frac{p_b+(1-2p_b)\left(1-\sum_j q_j\mu_j-\sum_j q_jP_{d,j}+2\sum_j q_j\mu_jP_{d,j}\right)}{p_b+(1-2p_b))\left(1-\sum_j q_j\mu_j-\sum_j q_jP_{f,j}+2\sum_j q_j\mu_jP_{f,j}\right)}$$

and

$$a^*_\ominus = \frac{1-p_b-(1-2p_b)\left(1-\sum_j q_j\mu_j-\sum_j q_jP_{d,j}+2\sum_j q_j\mu_jP_{d,j}\right)}{1-p_b-(1-2p_b)\left(1-\sum_j q_j\mu_j-\sum_j q_jP_{f,j}+2\sum_j q_j\mu_jP_{f,j}\right)}.$$

[6]Note that we consider the case where the Bayes cost ratio equals 1.

[7]This value is $\tilde{p}_i$ in (10) but the $\sim$ at the top has been dropped for notational simplicity.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2784358, IEEE Transactions on Signal Processing

If these data fusion components of Fig. 9 are from different crowds, one can go higher in the Bayesian hierarchical model and use the population parameters to determine the optimal fusion rule. Also, any machines using CV rules in the penultimate level of the hierarchical sociotechnical network can be regarded as a human agent with $\mathbf{q} = [1, 0, \ldots, 0]$ and perfect match value of $p = 1$. Such a generality can help us in constructing arbitrary-depth trees of sociotechnical decision making, where humans are modeled and the machines are optimized.

In the following, the benefit associated with the Bayesian hierarchical model is characterized. Consider the case when such a model of human decision fusion is ignored and are instead considered to be machines, then the optimal K for the K-out-of-M rule is given by

$$K^*_{sen} = \left\lceil \frac{\log \frac{P_0}{P_1} - M \log \frac{1-p_b-(1-2p_b)P_d}{1-p_b-(1-2p_b)P_f}}{\log \frac{(p_b+(1-2p_b)P_d)(1-p_b-(1-2p_b)P_f)}{(p_b+(1-2p_b)P_f)(1-p_b-(1-2p_b)P_d)}} \right\rceil . \quad (21)$$

From (20) and (21), we can observe that the basic difference between K^* and K^*_{sen} arises from the P_d and P_f of the intermediate decision fusion systems. If the intermediate decision fusion systems are machines, they have deterministic P_d and P_f, while the human decision fusion components modeled using the Bayesian hierarchical model have P_d and P_f that incorporate randomness. As we shall observe later in the paper, this incorporation of randomness into the optimal K improves system performance.

The error probability for fixed K is

$$\begin{aligned} P_e(K) = \\ P_0 \sum_{i=K}^{M} \binom{M}{i} \left(\tilde{P}_f\right)^i \left(1 - \tilde{P}_f\right)^{M-i} \\ + P_1 \sum_{i=0}^{K-1} \binom{M}{i} \left(\tilde{P}_d\right)^i \left(1 - \tilde{P}_d\right)^{M-i}, \quad (22) \end{aligned}$$

where

$$\begin{aligned} \tilde{P}_d = p_b + (1-2p_b) \\ \left(1 - \sum_j q_j \mu_j - \sum_j q_j P_{d,j} + 2 \sum_j q_j \mu_j P_{d,j}\right) \quad (23) \end{aligned}$$

and

$$\begin{aligned} \tilde{P}_f = p_b + (1-2p_b) \\ \left(1 - \sum_j q_j \mu_j - \sum_j q_j P_{f,j} + 2 \sum_j q_j \mu_j P_{f,j}\right). \quad (24) \end{aligned}$$

Therefore, the performance loss by ignoring the effect of humans in the system is due to the mismatched K value and is given by (25).

Fig. 10 shows the gain in performance by using the Bayesian hierarchical model of humans in comparison to assuming them to be machines, against prior probability for different values of N. The parameters used are $M = 5$, $P_d = [0.9, 0.8, 0.8]$ and $P_f = [0.1, 0.2, 0.3]$ for the three different rules, and the parameters $\mathbf{q}$, α, and β are the ones inferred from data and as listed in Table II. We plot the case when the channels are perfect ($p_b = 0$), to emphasize the gain associated with the models developed for human decision making in this paper. Fig. 10 clearly shows the high gain in performance by using the model developed in this paper. The gain in performance is highest for $N = 20$, i.e. when the number of sources for the decision fusion components is high. We observe some sudden jumps in performance gain around $P_0 = 0.1$ and $P_0 = 0.9$, and lack of performance improvement in the region around $P_0 = 0.5$. These regions are further explored for a simple case below.

For further insights, we consider the case when $R = 1$ in the following and only use the CV rule as a potential rule. In Fig. 11, the performance gain by using the Bayesian hierarchical model is plotted against different values of prior probability for this case. The parameters used are $M = 5$, $P_d = 0.9$, $P_f = 0.1$, $\alpha = 5$, and $\beta = 3$. As can be observed, by utilizing the knowledge of human decision fusion components in the system during system design, one can improve the performance by around 35% on average.

The sudden jump in performance gain around priors $P_0 = 0.1$ and $P_0 = 0.9$ is due to the chosen values of P_d and P_f and can be analytically determined using the expressions in (20) and (21). Also, note that the region around $P_0 = 0.5$ for which there is no performance improvement is due to the situation when the term dependent on the prior dominates the other terms in the expressions of K^* and K^*_{sen}, thereby resulting in equal values of K^* and K^*_{sen}. The width of this region where there is no performance gain depends on the values of α and β as we can see in Fig. 12. Here, $P_0 = 0.3$ is outside this region for $\beta \geq 1.5$ while it is within this region for $\beta < 1.5$. Similar observations can be made for different values of priors. This suggests that the performance gain with the Bayesian hierarchical model developed in this paper depends on the apriori information (P_0) about the task and the parameters of the crowd taking part in the task. As the crowd gets more unreliable (β increases), the proposed model can improve performance for a larger range of task prior probabilities.

VI. Discussion

In this paper, the human behavior in human-in-the-loop sociotechnical systems is studied. Specifically, the task of decision fusion has been considered. It was first observed that deterministic fusion rules, such as the CV rule, do not characterize human behavior, since data fusion by humans is not deterministic in nature. For a given set of data, deterministic rules give the same output at any time instant. On the other hand, the output changes for different humans and in some cases, for the same human at different time instants, as pointed out by Payne and Bettman [38]. This suggests the use of a randomized decision rule, which was the focus of the next part of the paper.

We developed hierarchical models which characterize this behavior. Due to the hierarchical nature, this model encompasses human variation observed at various levels: individual

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2784358, IEEE Transactions on Signal Processing

$$\Delta P_e = \begin{cases} \sum_{i=K^*_{sen}}^{K^*-1} \binom{M}{i} \left[P_0 \left(\tilde{P}_f\right)^i \left(1-\tilde{P}_f\right)^{M-i} - P_1 \left(\tilde{P}_d\right)^i \left(1-\tilde{P}_d\right)^{M-i} \right], & \text{if } K^* > K^*_{sen}, \\ \sum_{i=K^*}^{K^*_{sen}-1} \binom{M}{i} \left[P_1 \left(\tilde{P}_d\right)^i \left(1-\tilde{P}_d\right)^{M-i} - P_0 \left(\tilde{P}_f\right)^i \left(1-\hat{P}_f\right)^{M-i} \right], & \text{if } K^* < K^*_{sen} \end{cases} \tag{25}$$

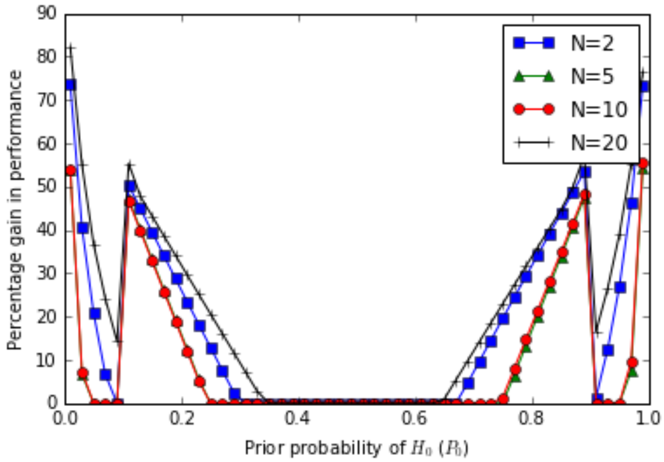

Fig. 10. Percentage improvement in system performance by using the Bayesian hierarchical model for system design with varying prior probability.

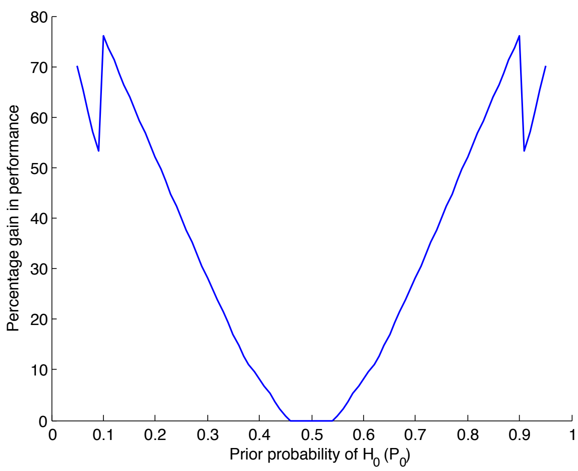

Fig. 11. Percentage improvement in system performance by using the Bayesian hierarchical model for system design with varying prior probability for a fixed CV rule.

level, crowd level, and population level. On an individual level, every human has a different bias which affects his/her decision fusion process. A crowd is a collection of people who have similar understanding due to cultural, societal, or other factors, and therefore, might have similar characteristics in performing tasks. On a population level, there are differences in societies, cultures, or demographics, which affect the decision fusion process. The effect of such models on the design of larger human-machine systems has been demonstrated. It was shown that there is a substantial improvement in performance when the human-behavior models are used for designing human-in-the-loop systems.

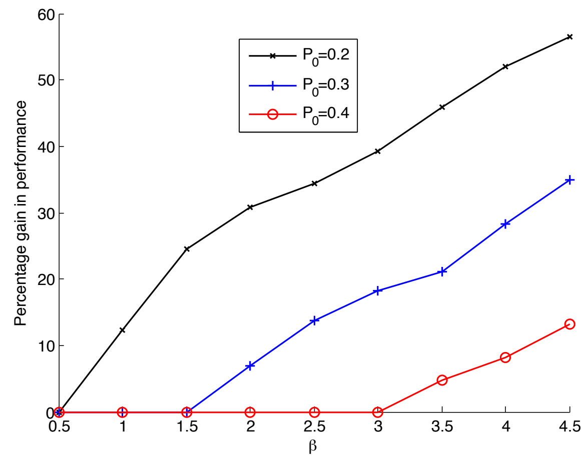

Fig. 12. Percentage improvement in system performance by using the Bayesian hierarchical model for system design with varying values of β and $\alpha = 0.5$.

This work demonstrates the benefits of the methodology involving the design of experiments to study human behavior, building statistical models that capture the essence of the observed human behavior, and using these models to optimize the design of large-scale human-machine systems. This methodology can be followed to model and understand other human user behavior. For example, data can be collected with a large number of sources (N) to verify some asymptotic approximations. In other words, this data can be used to verify the hypothesis that humans use heuristic decision rules when the amount of data is large. On similar lines, time-constrained tasks can be designed to verify if heuristic rules such as *pick-the-best* rule (MAD rule) work better under time-constrained situations. A psychological understanding of the observations might also provide insights towards comprehending complex human behavior. Computational social science data can also be used in lieu of psychology experiments used in this paper.

REFERENCES

[1] A. Vempaty, L. R. Varshney, G. J. Koop, A. H. Criss, and P. K. Varshney, "Decision fusion by people: Experiments, models, and sociotechnical system design," in *Proc. 3rd IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Dec. 2015, pp. 83–87.

[2] G. Baxter and I. Sommerville, "Socio-technical systems: From design methods to systems engineering," *Interaction Comput.*, vol. 23, no. 1, pp. 4–17, Jan. 2011.

[3] Y.-T. Chen, C. Caramanis, and S. Shakkottai, "On file sharing over a wireless social network," in *Proc. 2011 IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2011, pp. 249–253.

[4] P. Costa, C. Mascolo, M. Musolesi, and G. P. Picco, "Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 5, pp. 748–760, Jun. 2008.

[5] E. Stai, J. S. Baras, and S. Papavassiliou, "Social networks over wireless networks," in *Proc. 51st IEEE Conf. Decision Control*, Dec. 2012, pp. 2696–2703.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2784358, IEEE Transactions on Signal Processing

[6] P. Hui and J. Crowcroft, "How small labels create big improvements," in *Proc. Fifth IEEE Int. Conf. Perv. Comput. Commun. Workshops (PerComW'07)*, Mar. 2007, pp. 65–70.

[7] M. Musolesi and C. Mascolo, "Designing mobility models based on social network theory," *ACM SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 11, no. 3, pp. 59–70, Jul. 2007.

[8] P. K. Varshney, *Distributed Detection and Data Fusion*. New York: Springer-Verlag, 1996.

[9] V. V. Veeravalli and P. K. Varshney, "Distributed inference in wireless sensor networks," *Phil. Trans. R. Soc. A*, vol. 370, no. 1958, pp. 100–117, Jan. 2012.

[10] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I – Fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.

[11] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors: Part II – Advanced topics," *Proc. IEEE*, vol. 85, no. 1, pp. 64–79, Jan. 1997.

[12] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 1, pp. 98–101, Jan. 1986.

[13] M. Kam, Q. Zhu, and W. S. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 916–920, Jul. 1992.

[14] D. Bollier, *The Future of Work: What It Means for Individuals, Businesses, Markets and Governments*. Washington, DC: The Aspen Institute, 2011.

[15] D. Tapscott and A. D. Williams, *Wikinomics: How Mass Collaboration Changes Everything*. New York: Portfolio Penguin, 2006.

[16] ——, *Macrowikinomics: Rebooting Business and the World*. New York: Portfolio Penguin, 2010.

[17] Q. Li, A. Vempaty, L. R. Varshney, and P. K. Varshney, "Multi-object classification via crowdsourcing with a reject option," *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 1068–1081, Feb. 15 2017.

[18] A. Vempaty, L. R. Varshney, and P. K. Varshney, "Reliable crowdsourcing for multi-class labeling using coding theory," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 4, pp. 667–679, Aug. 2014.

[19] G. Gigerenzer and R. Selten, *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press, 2002.

[20] K. R. Varshney and L. R. Varshney, "Quantization of prior probabilities for hypothesis testing," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4553–4562, Oct. 2008.

[21] J. B. Rhim, L. R. Varshney, and V. K. Goyal, "Quantization of prior probabilities for collaborative distributed hypothesis testing," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4537–4550, Sep. 2012.

[22] T. Wimalajeewa and P. K. Varshney, "Collaborative human decision making with random local thresholds," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2975–2989, Jun. 2013.

[23] ——, "Asymptotic performance of categorical decision making with random thresholds," *IEEE Signal Process. Lett.*, vol. 21, no. 8, pp. 994–997, Aug. 2014.

[24] L. R. Varshney and K. R. Varshney, "Decision making with quantized priors leads to discrimination," *Proc. IEEE*, vol. 105, no. 2, pp. 241–255, Feb. 2017.

[25] J. B. Rhim and V. K. Goyal, "Distributed hypothesis testing with social learning and symmetric fusion," *IEEE Trans. Signal Process.*, vol. 62, no. 23, pp. 6298–6308, Dec. 2014.

[26] R. D. Sorkin, C. J. Hays, and R. West, "Signal-detection analysis of group decision making," *Psychol. Rev.*, vol. 108, no. 1, pp. 183–203, Jan. 2001.

[27] J. B. Soll and R. P. Larrick, "Strategies for revising judgment: How (and how well) people use others' opinion," *J. Exp. Psychol.*, vol. 35, no. 3, pp. 780–805, May 2009.

[28] T. L. Griffiths, C. Kemp, and J. B. Tenenbaum, "Bayesian models of cognition," in *The Cambridge Handbook of Computational Cognitive Modeling*, R. Sun, Ed. New York: Cambridge University Press, 2008, pp. 59–100.

[29] D. A. Balota, M. J. Yap, K. A. Hutchison, M. J. Cortese, B. Kessler, B. Loftis, J. H. Neely, D. L. Nelson, G. B. Simpson, and R. Treiman, "The English lexicon project," *Behav. Res. Methods*, vol. 39, no. 3, pp. 445–459, Aug. 2007.

[30] A. H. Criss and R. M. Shiffrin, "Interactions between study task, study time, and the low-frequency hit rate advantage in recognition memory." *J. Exp. Psychol.: Learn., Mem., Cogn.*, vol. 30, no. 4, pp. 778–786, Jul. 2004.

[31] P. Hemmer and A. H. Criss, "The shape of things to come: Evaluating word frequency as a continuous variable in recognition memory." *J. Exp. Psychol.: Learn., Mem., Cogn.*, vol. 39, no. 6, pp. 1947–1952, Nov. 2013.

[32] A. H. Criss, K. J. Malmberg, and R. M. Shiffrin, "Output interference in recognition memory," *J. Mem. Learn.*, vol. 64, no. 4, pp. 316–326, May 2011.

[33] G. J. Koop, A. H. Criss, and K. J. Malmberg, "The role of mnemonic processes in pure-target and pure-foil recognition memory," *Psychon. Bull. Rev.*, vol. 22, no. 2, pp. 509–516, Apr. 2015.

[34] J. Henrich, S. J. Heine, and A. Norenzayan, "Most people are not WEIRD," *Nature*, vol. 466, no. 7302, p. 29, Jul. 2010.

[35] T. A. Salthouse, "Age-related differences in basic cognitive processes: Implications for work," *Exp. Aging Res.*, vol. 20, no. 4, pp. 249–255, Oct.-Dec. 1994.

[36] J. Marschak and R. Radner, *Economic Theory of Teams*. New Haven: Yale University Press, 1972.

[37] J. B. Rhim and V. K. Goyal, "Social teaching: Being informative vs. being right in sequential decision making," in *Proc. 2013 IEEE Int. Symp. Inf. Theory*, Jul. 2013.

[38] J. W. Payne and J. R. Bettman, "Walking with the scarecrow: The information-processing approach to decision research," in *Blackwell Handbook of Judgment and Decision Making*, D. J. Koehler and N. Harvey, Eds. Oxford: Blackwell Publishing Ltd., 2004, pp. 110–132.

Aditya Vempaty (S'12) was born in Hyderabad, India, on August 3, 1989. He received the B.Tech. degree in electrical engineering from the Indian Institute of Technology (IIT), Kanpur, India, in 2011, with academic excellence awards for consecutive years, and the Ph.D. degree in electrical engineering from Syracuse University, Syracuse, NY, USA, in 2015.

He is a research staff member, at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA, where he was a postdoctoral researcher until 2016. He was a graduate research intern in the Data Systems Group at Intel Corporation, Santa Clara, CA, USA, in summer 2013 and a visiting student in the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, USA, in summer 2014. His research interests include human?machine inference networks, behavioral analytics, statistical signal processing, and network security.

Dr. Vempaty is a member of Phi Kappa Phi and Golden Key International Honor Society. He received the Syracuse University Graduate Fellowship Award in 2013 and the All University Doctoral Prize 2016 by Syracuse University for superior achievement in completed dissertations.

Lav R. Varshney (S'00–M'10–SM'15) received the B.S. degree (*magna cum laude*) in electrical and computer engineering with honors from Cornell University, Ithaca, New York, in 2004. He received the S.M., E.E., and Ph.D. degrees, all in electrical engineering and computer science, from the Massachusetts Institute of Technology, Cambridge, in 2006, 2008, and 2010, where his theses received the E. A. Guillemin Thesis Award and the J.-A. Kong Award Honorable Mention.

He is an assistant professor in the Department of Electrical and Computer Engineering, the Department of Computer Science (by courtesy), the Coordinated Science Laboratory, the Beckman Institute, and the Neuroscience Program at the University of Illinois at Urbana-Champaign. He is also leading curriculum initiatives for the new B.S. degree in Innovation, Leadership, and Engineering Entrepreneurship in the College of Engineering. During 2010–2013, he was a research staff member at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York. His research interests include information and coding theory; limits of nanoscale, human, and neural computing; human decision making and collective intelligence; and creativity.

Dr. Varshney is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi. He received the IBM Faculty Award in 2014 and was a Finalist for the Bell Labs Prize in 2014 and 2016. He and his students have won several best paper awards. His work appears in the anthology, *The Best Writing on Mathematics 2014* (Princeton University Press). He currently serves on the advisory board of the AI XPRIZE.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2784358, IEEE Transactions on Signal Processing

Gregory J. Koop received a B.A. in Psychology and History from Goshen College in 2006 and M.A. and Ph.D. degrees in psychology from Miami University, Oxford, OH, in 2010 and 2012, respectively.

From 2012-2014, he was a research fellow in the Memory Modeling Laboratory at Syracuse University. Since 2014, he has been an assistant professor in the Psychology Department at Eastern Mennonite University in Harrisonburg, VA. He is the author of 7 articles and book chapters in the domains of judgment and decision making and recognition memory. His research interests include evaluating computational models of preferential choice and recognition memory, the influence of reference points on risk preferences, and the application of ”nudges” to higher education.

Amy H. Criss received a B.A. in Psychology and Neuroscience from Miami University in 1997, and Ph.D. degree in Cognitive Psychology and Cognitive Science from Indian University in 2004.

Since 2007, she has been with Syracuse University, where he is currently a Chair and Full Professor of department of psychology. Her research involves describing how human memory operates within the framework of computational models. Her interests also include semantic knowledge and implicit memory.

Pramod K. Varshney (S'72–M'77–S'82–F'97) was born in Allahabad, India, in 1952. He received the B.S. degree (Hons.) in electrical engineering and computer science and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana?Champaign, IL, USA, in 1972, 1974, and 1976, respectively. From 1972 to 1976, he held teaching and research assistantships with the University of Illinois. Since 1976, he has been with Syracuse University, Syracuse, NY, USA, where he is currently a Distinguished Professor of electrical engineering and computer science and the Director of the Center for Advanced Systems and Engineering. He served as the Associate Chair of the department from 1993 to 1996. He is also an Adjunct Professor of radiology with Upstate Medical University, Syracuse. His current research interests are in distributed sensor networks and data fusion, detection and estimation theory, wireless communications, image processing, radar signal processing, and remote sensing. He has published extensively. He is the author of the book *Distributed Detection and Data Fusion* (New York: Springer-Verlag, 1997). He has served as a consultant to several major companies. Dr. Varshney was a James Scholar, a Bronze Tablet Senior, and a Fellow while at the University of Illinois. He is a member of Tau Beta Pi. He was elected to the grade of fellow of the IEEE in 1997 for his contributions in the area of distributed detection and data fusion. He was the President of International Society of Information Fusion during 2001. He was the recipient of the 1981 ASEE Dow Outstanding Young Faculty Award, the IEEE 2012 Judith A. Resnik Award, the degree of Doctor of Engineering honoris causa by Drexel University in 2014, and the ECE Distinguished Alumni Award from UIUC in 2015. He was the Guest Editor of the Special Issue on Data Fusion of the PROCEEDINGS OF THE IEEE in 1997. In 2000, he received the Third Millennium Medal from the IEEE and the Chancellor?s Citation for exceptional academic achievement at Syracuse University. He is on the Editorial Board of the *Journal on Advances in Information Fusion* and the *IEEE Signal Processing Magazine*.