2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)

Decision Fusion by People: Experiments, Models, and Sociotechnical System Design

Aditya Vempaty,*‖ Lav R. Varshney,† Gregory J. Koop,‡ Amy H. Criss,§ and Pramod K. Varshney¶
* IBM Thomas J. Watson Research Center
† Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign
‡ Department of Psychology, Eastern Mennonite University
§ Department of Psychology, Syracuse University
¶ Department of Electrical Engineering and Computer Science, Syracuse University

***Abstract*—People and machines perform tasks differently. Building optimal systems that include people and machines, requires understanding their respective behavioral properties. The task of decision fusion is considered and the performance of people is compared to the optimal fusion rule. Our behavioral experiments demonstrate that people perform decision fusion in a stochastic manner dependent on various factors, whereas optimal rule is deterministic. A Bayesian hierarchical model is developed to characterize the observed human behavior. This model captures the differences observed in people at individual level, crowd level, and population level. The implications of such a model on developing large-scale human-machine systems are presented by developing optimal decision fusion trees with both human and machine agents.**

***Index Terms*—human decision-making, decision fusion, Bayesian hierarchical modeling, sociotechnical systems**

I. Introduction

Decision-making is the process of selecting an alternative among multiple choices based on the collected evidence. Such processes are highly prevalent in our daily life. 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.

In the signal processing literature, such problems have been extensively studied when all the decision makers are machines/sensors [1]–[4] and optimal decision rules at both LDM and GDM have been designed under various assumptions [1], [5], [6]. When the GDM is an optimized fusion rule but the LDMs are humans, the above framework addresses the paradigm of crowdsourcing for distributed inference tasks [7]–[9]. In such systems, one can analyze the system performance and design simple easy-to-perform tasks to improve the overall performance of the system [10].

To engineer systems where the GDM is also a human, it is of interest to understand how people fuse decisions. This work studies decision fusion by humans via behavioral experiments and demonstrates differences from optimal approaches. Based on experimental results, we develop a particular bounded rationality model (cf. [11]).

The problem of fusing multiple human decisions has been investigated in different contexts in the psychology literature (see [12], [13] and references therein). Such a framework is very similar to the problems in social choice theory, and voting. Based on these prior works and also our new experimental data, it can be inferred that human behavior is non-deterministic in general. Therefore, Bayesian hierarchical modeling approaches [14] are adopted to describe such behavior. Due to the hierarchical nature, this model encompasses

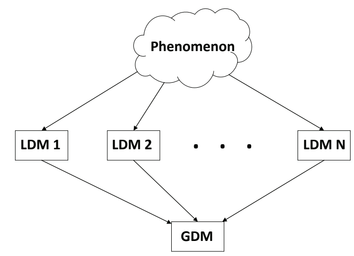

Fig. 1. System model consisting of LDMs and a GDM.

human variation observed at various levels: individual 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.

Making use of our Bayesian hierarchical model of human behavior, we develop optimal decision fusion trees with both human and machine agents. In particular, we incorporate the randomness associated with human behavior into the design of fusion rules and show the improvement in performance by using the fusion rule developed in this paper.

The remainder of the paper is organized as follows: In Sec. II, we describe psychology experiments to understand human decision fusion, especially in comparison to optimal fusion rules. After establishing that the existing decision fusion models of machines cannot explain the human behavior, in Sec. III, we build Bayesian hierarchical models to explain the observed behavior. In Sec. IV, we discuss its implications by demonstrating its effect on the design of large-scale sociotechnical systems. We conclude the paper in Sec. V.

II. Experiments and data analysis

Sec. II-A describes the experiments designed to study decision fusion by humans. The collected data is analyzed in Sec. II-B and the performance of humans is compared with that of optimal rules.

A. Experimental design

To understand the decision fusion behavior in humans, experiments replicating the process of Fig. 1 were designed. Human subjects consisting of undergraduate students of Syracuse University were enrolled for this task. The experiment consisted of data collection in two stages: the first stage models the local decision-making and the second stage models the data fusion aspect. The experiment is that of a memory-based task and is described as follows. Consider a target set of 100 words $\mathcal{D}$ and a distinct set of 100 distractor words $\mathcal{N}$,

‖ Work performed while with Syracuse University.

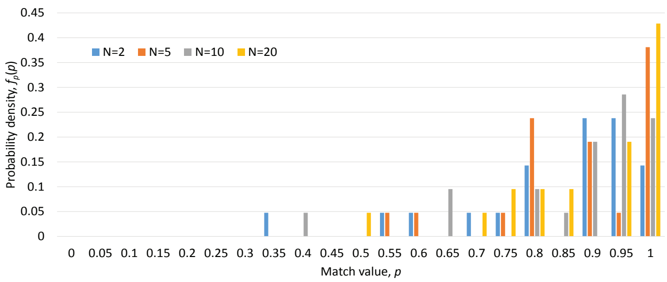

Fig. 2. Distribution of subjects' match value between the human decision and the CV rule's decision.

with $\mathcal{S} = \mathcal{D} \cup \mathcal{N}$. For the first stage, human subjects (called sources) took part in a recognition task where they first memorized $\mathcal{D}$, then performed a local decision for each $s \in \mathcal{S}$ as to whether $s \in \mathcal{D}$ or $s \in \mathcal{N}$. In the second stage, a new set of human subjects had to decide whether the word was present in the original database $\mathcal{D}$ by using decisions from the sources. These human subjects of second stage replicate the role of a GDM (Fig. 1). Note that these decision makers of second stage have no direct access to the database; their only source of information is from the sources. Local decisions from a variable number of sources (N) were presented to these subjects. This value N was either 2, 5, 10, or 20. The subjects were also presented with the sources' reliabilities and bias values. These values play the role of probability of detection and false alarm used in signal processing literature. Each dataset 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 and reliabilities, and the fused decision reported by GDM.

B. Data analysis

This section presents a summary of the analyzed data. First, the optimal decision fusion rule [5] is presented for comparison.

Optimal fusion rule: When the sources' reliabilities are known, optimal decision fusion is achieved by the Chair-Varshney (CV) rule [5]. Represent the "Yes/No" decisions of ith LDM as

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

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

$$u_0 = \begin{cases} +1, & \text{if } a_0 + \sum_{i=1}^{N} a_i u_i > 0, \\ -1, & \text{otherwise}, \end{cases} \quad (2)$$

where $a_0 = \log \frac{P_1}{1-P_1}$,

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

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

How efficient are people?: To determine how efficient humans are at fusing decisions, the final decisions by 21 human subjects are compared with the decision from the Chair-Varshney rule.[1] 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 match the optimal rule around 80–90% of the time. The closeness with the optimal fusion rule increases with an increase in N from 2 to 20. By defining the *match value* of a subject as the fraction of times his/her decision matches the decision of the CV rule with the same input data, individual participant's performance is compared. Although there is 80–90% match overall, closer examination shows that the individual match value has a lot of variation across subjects. For example, when $N = 5$, while one participant had a low match value of 0.54, another participant had a high match value of 0.98. Fig. 2 shows the distribution of the match value between the human's decision and the CV rule's decision for different values of N. Therefore, a single decision fusion rule (such as the CV rule) cannot capture the human behavior. Next, we develop a model to represent the observed human behavior.

[1]Note that in our setup, $P_1 = 0.5$, implying $a_0 = 0$.

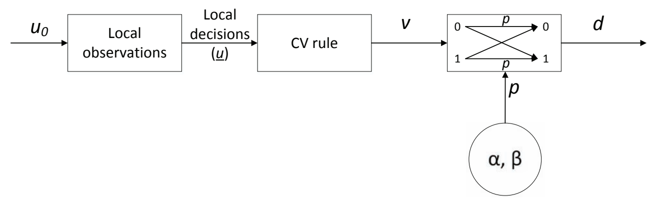

Fig. 3. 2-step model where the first step determines a deterministic decision using CV rule and the second step models the randomness of human decision-making. Here α and β are hyperparameters that capture the randomness in match value.

III. Bayesian hierarchical model

In this section, a Bayesian hierarchical model is developed which characterizes the human behavior at fusing multiple decisions.

A. Description of model

The phenomenon of different match values across individual participants can be represented as a random variable following a distribution as shown in Fig. 2. 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 optimal fusion rule (CV rule).
- The next step is a randomization step, where a match value p is sampled from a distribution $f_p(\cdot)$.

This distribution $f_p(\cdot)$ is determined by fitting a model to experimental data in Fig. 2. The final decision is now given by:

$$d = \begin{cases} v, & \text{with probability } p, \\ 1 - v, & \text{with probability } 1 - p. \end{cases} \quad (3)$$

Due to the limited number of data points, a bootstrap model is used for data fitting, where $n = 15$ data points among the total $T = 21$

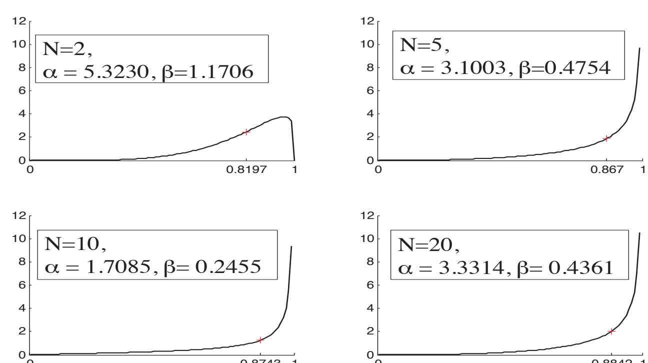

Fig. 4. Distribution $f_p(\cdot)$ of match value p, based on data fitting. The mean value is also highlighted.

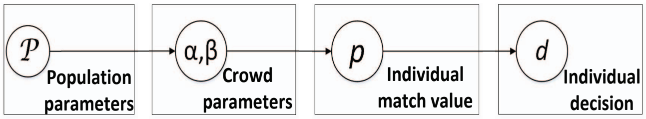

Fig. 5. Bayesian hierarchical model of decision fusion by humans.

data points are randomly selected for which a Beta distribution with parameters α and β are fit. This process is repeated $N_{mc} = 1000$ times. If α_j and β_j represent the parameters from the jth trial, the final parameters are decided by taking an average of these parameters. The results are shown in Fig. 4. An interesting observations is that the distribution $f_p(\cdot)$ shifts to the right and the mean increases with increase in N.

Clearly the exact values of α and β are themselves dependent 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 α and β, or in other words, the distribution $f_p(\cdot)$ itself is dependent on the underlying crowd chosen for the task. Different crowds would have different values of 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 α and β can be modeled as random variables sampled from a distribution with parameters $\mathcal{P}$ (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. 5.

IV. Optimal Design of Sociotechnical Systems

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 their effect on the crowd parameters α and β. For example, one might get different results from online participants, such as Turkers from Amazon Mechanical Turk[2], as compared to a group of college students [15]. 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 [1], [16] where each agent is not a single cognitive agent, but rather a human-based decision fusion system (Fig. 6). Also, cognitive agents in such systems may be drawn from a specialized sub-population.

Here we consider designing sociotechnical systems with machines and with humans, as modeled through our hierarchical Bayesian framework. Consider a system like Fig. 6 where multiple levels of decision makers are present in the system with human decision makers fusing data from multiple subordinate agents (humans or sensors) before sending their fused observations to a final fusion center. If these last level agents were sensors/machines rather than humans, one can use the optimal fusion rule to fuse the data [5]. 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, one needs to use the Bayesian hierarchical model of human decision fusers to design the fusion rule at the FC. [3]

Considering the Bayesian formulation, the optimal fusion rule at the FC are developed by adopting a methodology similar to [5]. Let

[2]https://www.mturk.com/

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

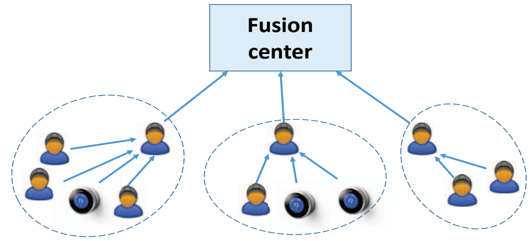

Fig. 6. Hierarchical system consisting of human decision fusion components.

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 FC receives decisions from N_c human decision fusion components. We represent their decisions by $d_i \in \{-1, +1\}$, where $d_i = -1(+1)$, if the ith component's decision is $H_0(H_1)$. The FC makes the final decision $d_0 = f(d_1, \ldots, d_{N_c})$ using the N_c 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. 5).

The optimal decision rule that minimizes the probability of error at the FC is given by the following likelihood ratio test

$$\frac{P(d_1, \ldots, d_{N_c}|H_1)}{P(d_1, \ldots, d_{N_c}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0}{P_1}, \tag{4}$$

or equivalently,

$$\log \frac{P(H_1|d_1, \ldots, d_{N_c})}{P(H_0|d_1, \ldots, d_{N_c})} \underset{H_0}{\overset{H_1}{\gtrless}} 0. \tag{5}$$

This optimal fusion rule can be written as

$$\log \frac{P_1}{P_0} + \sum_{\mathcal{S}_\oplus} \log \frac{P(d_i = +1|H_1)}{P(d_i = +1|H_0)} + \sum_{\mathcal{S}_\ominus} \log \frac{P(d_i = -1|H_1)}{P(d_i = -1|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} 0, \tag{6}$$

where $\mathcal{S}_\oplus(\mathcal{S}_\ominus)$ is the set of all components that reported a decision $d_i = +1(-1)$.

The terms in (6) can be further simplified as

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

where $d_{i,CV} \in \{-1, +1\}$ is the decision that the ith human fusion center would make using the optimal CV rule, p_i is the match value of the ith human, and $P_{d,i} \triangleq P(d_{i,CV} = +1|H_1)$. 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 $P_{f,i} \triangleq P(d_{i,CV} = +1|H_0)$.

This simplifies the optimal fusion rule (6) as

$$\log \frac{P_1}{P_0} + \sum_{\mathcal{S}_\oplus} \log \frac{1 - p_i - P_{d,i} + 2p_i P_{d,i}}{1 - p_i - P_{f,i} + 2p_i P_{f,i}} + \sum_{\mathcal{S}_\ominus} \log \frac{p_i + P_{d,i} - 2p_i P_{d,i}}{p_i + P_{f,i} - 2p_i P_{f,i}} \underset{H_0}{\overset{H_1}{\gtrless}} 0. \tag{7}$$

Note that the above expression requires the knowledge of individual match values. When this knowledge is not available, but the crowd parameters α and β are known (refer to Fig. 5), it is not very difficult to see that the only change in the optimal fusion rule will be to replace p_i with $E[p_i] = \frac{\alpha}{\alpha+\beta}$. Therefore, when all the decision

$$\Delta P_e = \begin{cases} \sum_{i=K^*_{sen}}^{K^*-1} \binom{N_c}{i} \left[P_0 \left(\frac{\beta+(\alpha-\beta)P_f}{\alpha+\beta}\right)^i \left(\frac{\alpha-(\alpha-\beta)P_f}{\alpha+\beta}\right)^{N_c-i} - P_1 \left(\frac{\beta+(\alpha-\beta)P_d}{\alpha+\beta}\right)^i \left(\frac{\alpha-(\alpha-\beta)P_d}{\alpha+\beta}\right)^{N_c-i} \right], & \text{if } K^* > K^*_{sen}, \\ \sum_{i=K^*}^{K^*_{sen}-1} \binom{N_c}{i} \left[P_1 \left(\frac{\beta+(\alpha-\beta)P_d}{\alpha+\beta}\right)^i \left(\frac{\alpha-(\alpha-\beta)P_d}{\alpha+\beta}\right)^{N_c-i} - P_0 \left(\frac{\beta+(\alpha-\beta)P_f}{\alpha+\beta}\right)^i \left(\frac{\alpha-(\alpha-\beta)P_f}{\alpha+\beta}\right)^{N_c-i} \right], & \text{if } K^* < K^*_{sen} \end{cases} \quad (8)$$

fusion components are identical (same number of sources, identically distributed sources, etc.), then the optimal fusion rule becomes a K out of N rule. The optimal K^* is easy to derive and is given by

$$K^* = \left\lceil \frac{\log \frac{P_0}{P_1} - N_c \log a^*_{\ominus}}{\log \frac{a^*_{\oplus}}{a^*_{\ominus}}} \right\rceil, \quad (9)$$

where

$$a^*_{\oplus} = \frac{1 - \frac{\alpha}{\alpha+\beta} - P_d + 2\frac{\alpha}{\alpha+\beta}P_d}{1 - \frac{\alpha}{\alpha+\beta} - P_f + 2\frac{\alpha}{\alpha+\beta}P_f} = \frac{\beta + (\alpha-\beta)P_d}{\beta + (\alpha-\beta)P_f}$$

and

$$a^*_{\ominus} = \frac{\frac{\alpha}{\alpha+\beta} + P_d - 2\frac{\alpha}{\alpha+\beta}P_d}{\frac{\alpha}{\alpha+\beta} + P_f - 2\frac{\alpha}{\alpha+\beta}P_f} = \frac{\alpha + (\beta-\alpha)P_d}{\alpha + (\beta-\alpha)P_f}.$$

If these data fusion components of Fig. 6 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 system can be regarded as a human agent with a perfect match value of 1. Such a generality can help us in potentially 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, then the optimal K^*_{sen} for the K out of N rule is given by

$$K^*_{sen} = \left\lceil \frac{\log \frac{P_0}{P_1} - N_c \log \frac{1-P_d}{1-P_f}}{\log \frac{P_d(1-P_f)}{P_f(1-P_d)}} \right\rceil. \quad (10)$$

The error probability for fixed K is

$$P_e(K) = P_0 \sum_{i=K}^{N_c} \binom{N_c}{i} \left(\frac{\beta+(\alpha-\beta)P_f}{\alpha+\beta}\right)^i \left(\frac{\alpha-(\alpha-\beta)P_f}{\alpha+\beta}\right)^{N_c-i} + P_1 \sum_{i=0}^{K-1} \binom{N_c}{i} \left(\frac{\beta+(\alpha-\beta)P_d}{\alpha+\beta}\right)^i \left(\frac{\alpha-(\alpha-\beta)P_d}{\alpha+\beta}\right)^{N_c-i}.$$

Therefore, the performance loss by ignoring the effect of humans in the system is due to the mismatched K value and is given by (8).

In Fig. 7, the performance gain by using the Bayesian hierarchical model is plotted against different values of prior probability. The parameters used are $N_c = 5$, $P_d = 0.9$, $P_f = 0.1$, $\alpha = 5$, 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 an 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 (9) and (10). 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. 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. 8, as $P_0 = 0.3$ 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.

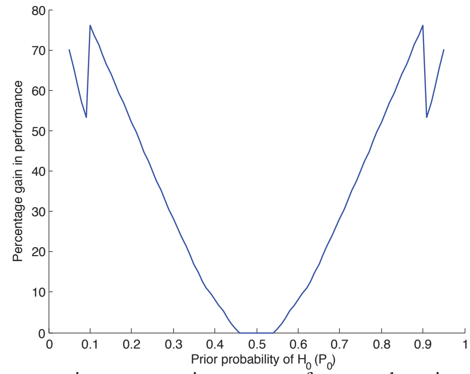

Fig. 7. Percentage improvement in system performance by using the Bayesian hierarchical model for system design with varying prior probability.

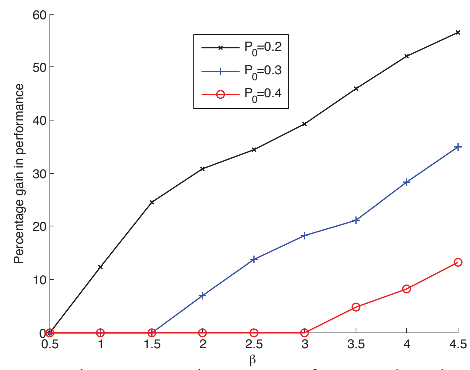

Fig. 8. Percentage improvement in system performance by using the Bayesian hierarchical model for system design with varying values of β and $\alpha = 0.5$.

V. Discussion

In this paper, the decision fusion problem has been considered. It was first observed for a system involving humans that a deterministic optimal fusion rule, such as the Chair-Varshney rule, does not characterize the human behavior, since data fusion by humans is not deterministic in nature. For a given set of data, the optimal deterministic rule gives 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 instant, as pointed by Payne and Bettman in [17]. This suggests the use of a randomized decision rule, which was the focus of the next part of the paper. Hierarchical models have been developed which characterize this behavior. The effect of such models on the design of larger human-machine systems has been demonstrated.

Currently, the data is being analyzed to identify the individual cases when the decisions of humans do not match the CV rule's decision. A psychological understanding of these particular cases can help us in comprehending this complex phenomenon. Time-constrained tasks are being considered, to verify if heuristic rules such as *pick-the-best* rule would work under such time-constrained situations.

REFERENCES

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

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

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

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

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

[10] 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.

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

[12] 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.

[13] 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.

[14] 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.

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

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

[17] 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, UK: Blackwell Publishing Ltd., 2004, pp. 110–132.