Fusion of soft information using TBM

Dafni Stampouli
Innovation Works
EADS UK
Newport, Wales, UK
Dafni.Stampouli@eads.com

Martin Brown
Innovation Works
EADS UK
Newport, Wales, UK
Martin.Brown@eads.com

Gavin Powell
Innovation Works
EADS UK
Newport, Wales, UK
Gavin.Powell@eads.com

Abstract* – *Intelligence operations are dependent on humans for data gathering and processing. Measurements made by human observation are subjective and incorporate biases. Systematic processing of such data is not as straightforward as with electronic sensor data. This paper presents work in progress towards developing a system to fuse inconsistent information and minimising observation errors. A case-study is presented, relating the techniques to the application area of civilian intelligence systems. Within our system, the Transferable Belief Model (TBM) was used to fuse soft information from observers, and combine in both the discrete and continuous spaces.

Keywords: Soft Data Fusion, HUMINT information, Transferable Belief Model, Civilian Intelligence.

1 Introduction

Police operations, and in particular intelligence analysis, often require to search, link and comprehend information relating to different entities (such as people, vehicles, locations) involved in the incident under investigation. Furthermore, the analyst is trying to comprehend the evidence, determine links among entities as well as identify leads. Previous work [1] introduced research into the development of the next generation of Command and Control, and Intelligence Analysis computer systems. The proposed framework is being targeted at the UK civilian authorities such as Police, Fire and Rescue, and Ambulance Services, but is also applicable to military operations. The proposed system is designed to combine data and information from diverse sources in order to provide situation awareness and improve decision making. The proposed system receives information from a variety of sources including traditional sensors (IMINT - IMagery INTelligence and COMINT - COMmunications INTelligence) and human reports (HUMINT - HUMan INTelligence). Human-based observation is an important data source in civilian intelligence systems. Furthermore, the collection of raw data might include both openly available sources such as the Internet and closed sources such as national databases, accessed by humans or via technical means. Soft data, as opposed to hard data, refers to data extracted from non-electronic sensors, such as intelligence reports, and open source material available on the internet. Therefore, human sensors may include Police reports obtained during investigation, or observation, reports from the public (both victims and witnesses), as well as individual informers. Such data by its nature is the result of a subjective assessment. The imprecision of, and variability between, human subjects has led Police in the UK to form the National Intelligence Model (NIM) in order to standardise information gathering and processing and grade soft information according to its source, before it is further processed.

This paper, initially discusses issues associated with using soft data from human intelligence and then reports on our research addressing some of the problems of dealing with such data. We describe the Transferable Belief Model (TBM) which we use to classify conflicting data and fuse together contradictory witness statements.

The paper is organised in seven sections, Section 2 discusses issues associated with soft data. Section 3 looks at the categorisation of attribute values, and Section 4 reviews the methodology and the algorithms we used to address the identified problems. Section 5 presents a case study to demonstrate the concepts and methodologies. Section 6 discusses the results and presents their evaluation. While the paper closes with some concluding remarks and discussion of future work in Section 7.

2 Issues with soft data

Emergency services, and in particular policing operations, benefit from the utilisation of humans in order to collect, report and make sense of information. The primary advantage of using human sensors is their ability to notice and report on relationships and inferences which are not provided by traditional sensors, which only measure features and attributes of certain entities. There are however several issues with the employment of human sensors as reported by Hall et.al. [2] as discussed in this section.

In the specific application area of civilian intelligence soft data sources include investigational reports, and observational statements taken from the public (both victims and witnesses). Increasingly, valuable data is obtained from open source material available on the internet. Initially there is a need to quantify the reliability and uncertainty of soft data. Furthermore, there are issues on how to extract information and knowledge from the reports. Finally there are issues in the quality, subjectivity and even format, of the data. In this section some of these quality issues are reviewed in greater detail, along with methods on how they can be addressed.

2.1 Elicitation

Elicitation refers to the process of obtaining the data in the first place. Some data may be unsolicited and freely available such as those on the internet. Others may be solicited and obtained through techniques such as surveys, questionnaires, debriefing and interviewing. The choice of technique and the experience of the agent eliciting the data have a huge bearing on its qualities and substance. The matter of elicitation of soft-data is discussed extensively in [2] and [3].

2.2 Source characterization and grading

One may consider one calibrated electronic source to be the equal of another of the same type, but soft data sources are not all equal. One may be considered more reliable or accurate than another. For example, we may give more credence to the description of a suspect taken from a trained police-officer than that taken from an ordinary member of the public. In previous work [1] we recognised the importance of grading the quality of data sources involved in the fusion process. We proposed an extension to the credibility model of [4] to fit with the UK National Intelligence Model (NIM) [5]. The NIM suggested 5x5x5 standard allows the recording of the reliability, origin, competency, and opportunity of the source.

2.3 Imprecision and omission

Humans are not precise measuring devices. Consider the scenario of a witness reporting the height of the suspected perpetrator of a crime. They may estimate the suspect to be 1.77 m when in fact they are 1.72 m. This imprecision is natural and not unexpected. They may give a range of values such as between 1.7 m and 1.85m, or say the suspect was of ‘average height’, or ‘slightly taller than me’. This becomes an issue since humans are required to portray very specific information with subjective descriptions.

The very act of recording reported soft data may introduce inaccuracies. This could be the result of transcription errors (‘typos’), misunderstandings or any one of a multitude of differences in cultural, language, and regional variations. For example, different regional spellings – ’Brown’ and ‘Browne’, different language spellings – ‘Richard’ and ‘Ricardo’. Handling inaccuracies such as these is a widespread problem. In the allied field of ‘Record Linkage’ the problem is known variously as data cleaning, data cleansing and data standardisation. For string based data values there are established means to recognise, correct or otherwise deal with such inaccuracies.

Omissions in soft data may result from the human result not taking place (eg ‘I didn’t see’) or from the inability to interpret the assessment (eg ‘I don’t know what make of car it was’). This difference may or may not be of significance, depending on the remaining evidence.

2.4 Biases

Measures resulting from human assessments are not only subject to imprecision and omission but also to subjective bias. Biases may reflect some quasi-permanent characteristic of the observer e.g. their personal, cultural, or national beliefs, values, prejudices or other characteristics. Assessments repeated by the same observer may exhibit traits that would allow a partial compensation for such quasi-permanent biases. Biases could also however be due to transient or otherwise temporary effects and more than likely impossible to compensate for. Biased assessments are more than likely subconscious and unintentional but conscious intentional biases cannot be ruled out. Some references discussing biases are [11] and [12].

2.5 Differences and conflicts

In all but the most simple of scenarios, there is unlikely to be complete, unequivocal agreement in the measures from several human observers. Instead there will likely be several different results. For example, one witness may report a suspect’s height to be 1.8m, another “between 1.8 and 1.9m and another “quite tall”. Despite the differences in the styles of the responses and values given, we might consider there to be some agreement and consistency across these responses. In contrast, conflicts occur when we consider there to be no agreement or consistency across responses. For example, one witness reports the colour of a car to be black and another witness blue. Differences and conflicts in data from human assessments can be handled by a data fusion process and this is discussed in Section 5.

3 Attribute-value categorisation

As can be seen from the discussion above, the elicitation and processing of soft data presents a number of challenges – the data is generally subject to biases, open to interpretation and misinterpretation, and often unstructured. In this paper the scenario of a guessing a person’s height and the colour of a vehicle are used to illustrate the difficulties of processing soft data and explain the methods we have developed to solve those issues. In Section 5 we will show that asking a group of people to characterise those simple attributes is not a trivial task, and may result in large inter- and intra-observer variabilities. We also focus on the fact that the usability of the description used depends greatly on the format that is used to represent the data. And in Section 6, we will prove that a single representation is not adequate to describe even a simple attribute. But first we list the categories that are used to describe different attribute values.

As discussed in Section 1, human beings are imprecise and biased measuring devices and the data they produce is typically unstructured. The results of a human-measurement of an identifying attribute – the attributes value – may vary widely across observers. To simplify the discussions, the whole spectrum of possible attribute values is restricted so that the values can be categorised into a small handful of types (omissions are ignored in these discussions). We follow the measurement scale types first proposed by Stevens [13]:

- **Nominal** The values are simply labels, and no order or measurement can be implied from these labels. Example nominal values in the assessment of a vehicle's colour are 'green' and 'black'.
- **Ordinal** Values have a natural order or rank in some way but no explicit measurement can be implied. Example ordinal values for a person's height are 'tall', 'average height', 'short'. Furthermore there is no concept of distance between one ordinal value and the next, nor a linear scale spanning all values.
- **Ratio** Represent continuous values, in some finite or infinite interval, which can be ordered and where there is a natural distance measure. For example, the value 1.7m for a person's height.
- **Range** In addition to the above measurement scale types proposed by Stevens, we permit the result to be a range of values. For the assessment of a person's height this permits responses like 'average' to 'tall', and 'between 1.8 and 1.9 metres'. These are natural responses for humans to give.

4 Methodology

The TBM is an extension of Dempster-Shafer theory; it is a two-level model – where beliefs held by an agent are represented at the *credal* level by belief functions, whereas decision making is based on probability functions at the *pignistic* level. The TBM has been primarily developed for discrete frames of discernment [6]. Later on, Smets [7] defined belief functions on continuous frames, where belief masses generalise into belief densities, by assuming that beliefs are quantified by belief functions. Further information about belief functions and the TBM can be found in Shafer [8] and Smets [9]. Caron et al [10] provided the explicit formulation of the least committed *Basic Belief Density (bbd)* of an *n*-dimensional Gaussian *Probability Density Function* (*pdf*), which we use to assign belief based on multidimensional continuous prior probabilities. Powell et.al [14] has used continuous priors to classify targets using speed characteristics. This paper presents a novel application of the continuous multidimensional TBM as described by Caron et. al. [10] in fusing and classifying soft information, in this case the reported observations of the colour of a vehicle and the height of a person.

So for example, in the case of colour, we apply it to a colour classification problem and to an information fusion problem. We map from the continuous Red, Green, Blue (RGB) colour space to the discrete space modelled by descriptive words. These words describe colours and by using a *Probability Density Function* (3-dimensional Gaussian) for each word we can map between RGB space and the words used to describe colours. First we classify each RGB triple in the continuous space, obtained from each witness statement, to a descriptive word, prior to fusion. Secondly we combine the discrete opinions of all the observers, that state a percentage of belief to each of the possible colour names, using the discrete TBM. Then the results from the two processes are fused together. This is a novel application of the TBM fusing between continuous and discrete colour spaces.

4.1 Transferable Belief Model

This section provides a brief overview of the theories of the TBM in both the discrete and continuous domains.

In the case of the discrete TBM, the frame of discernment, Ω, contains all the possible outcomes, $\Omega = \{x_1, x_2, \ldots x_n\}$. An agent's beliefs are quantified using a *basic belief assignment* (*bba*):

$$m : 2^{\Omega} \to [0,1] \text{ with } \sum_{A \subseteq \Omega} m(A) = 1.$$

Each *basic belief mass* (*bbm*), (A), is interpreted as the amount of support that an agent gives $A \subseteq \Omega$. The subsets $A \subseteq \Omega$ that have mass $m(A) > 0$ are called focal sets. Consequently, the belief and plausibility are defined, respectively, as:

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \quad (1)$$

$$pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \quad (2)$$

For the continuous case, a *bba* will become a *basic belief density* (bbd) [7]. Using the least committed *bbd* theory, we can construct the relevant underlying conditional *bbd*s for the required *pdf*, which will give the mass, probability and belief. By using the Generalised Bayesian Theorem (GBT), a conditional *bbm*, $m(A \mid x)$, for $A \in 2^{\Omega}$, and $x \in \Re^n$, can be calculated for a given class using:

$$m(A \mid x) = \prod_{\omega_i \in A} pl(x \mid \omega_i) \prod_{\omega_i \in \bar{A}} \left[1 - pl(x \mid \omega_i)\right]$$

where *pl(.|.)* is the least committed conditional plausibility.

Thereafter, the pignistic transform can used to provide us with the probabilities of every class, according to:

$$BetP\{\omega_i \mid x\} = \sum_{A:\omega_i \in A} \frac{1}{|A|} \frac{m(A \mid x)}{[1 - m(\emptyset \mid x)]}$$

This can then be used to make decisions.

More details and the working of the methods can be found on [7] and [10].

5 Case Study

In this section, a simple case study is used to demonstrate the issues discussed previously and describe the work that attempts to address them. The questionnaire used in our case study consisted of two parts: a) the observers were asked about characteristics of objects that were known and fixed, in this case the colour of a vehicle, and b) about characteristics that were to be estimated, in this case the height of a person. In the processing of the observer's responses, we quantify the variabilities in descriptions of human observers, in terms of accuracy and precision. The inter-observer variability and intra-observer variability were calculated. The answers were then fused, and the results were evaluated in terms of accuracy. The difference between fusion results and observer variabilities are also calculated to assess the performance of the fusion algorithm.

5.1 Vehicle Colour

The set of fourteen observers were shown pictures of two vehicles. The experiment used pictures, to test the variabilities in factual information, eliminating the variations on the recall ability, and memory bias. For each vehicle shown in a picture, they were asked to state its colour, with questions set in four different forms:

1. Free text. The observers were asked to state the colour they see.
2. Choose an answer from a list of colour names, where observers were allowed one answer.
3. Choose an answer from the same list of colour names where observers were allowed more than one answers but were expected to give a percentage indicating how strong they feel for each answer.
4. Choose an answer from a palette of colours similar to those indicated on the list (one answer allowed).

Therefore, the observers were asked to answer the same question in different forms, as discussed in Section 3, in order to corroborate their answers and reduce biases. This has proven to be a very useful step, as it eliminated random errors in the answers. Further explanation and examples will follow in later sections.

5.2 Human Height

The same set of observers were questioned about the height of two subjects. For each subject, the observers were asked to state their height, with questions set in two different forms:

1. Give a numerical estimate of height.
2. Choose one from five height categories:

$\Omega = \{very\ short, short, medium\ height, tall, very\ tall\}$

Then the same group of observers were asked to repeat the exercise, in order to quantify intra-observer variability. The second time the observers were asked to state their own height and the height ranges that they would give for each of the five categories. The two subjects were of average height (1.72m and 1.77m).We avoided using subjects that were very tall or very short, in order to have greater diversity in the answers.

5.3 Attribute fusion

For each entity in the experiment, the answers were fused to eliminate the variabilities and accurately classify the entity. This resulted in a significant increase in accuracy and precision. The fusion helps to resolve conflict in responses and provides a more consistent result.

6 Results and Evaluation

The efficacy of an observation can be assessed in terms of two factors: precision, and accuracy. Precision (also called reliability) describes how closely different observations of the same thing agree with each other, therefore tests repeatability of observation taking into account the influence of all subjective actions that are required to produce the result. Standard deviation and the coefficient of variation are measures of precision. Accuracy (also called validity) describes how closely the observation result agrees with the true or most probable value. Note that a high accuracy (i.e. high validity) means that the mean of repeated results is close to the true value, while a low precision (i.e. low reliability) means that the results are scattered. A low accuracy will generate more difficulties when interpreting results than a low precision. Intra-observer variability is assessed when the same observer estimates the same object in the same scene twice. When two or more observers estimate the same object in the same scene once, the inter-observer variability can be assessed. Evaluation of the observer related variabilities is an issue of critical importance, as the evidence they provide will, in the application area of civilian intelligence, be used to support an investigation, and might end up convicting a person in court. In other applications, such as medical diagnosis, the issue of observer variabilities has been researched extensively. To the authors' best knowledge analysis of the variabilities of observation has not been applied to intelligence operations. It is expected that the variability in those would be considerable. Due to this we designed a simple example in order to minimise biases and obtain highly correlated results. The findings were enlightening, as discussed in the next sections.

6.1 Vehicle Colour Estimation

In the first part of the questionnaire the 14 observers were shown a picture of two vehicles and were asked to use free text to describe their colour. For the first vehicle, denoted ‘Car 1’, the agreement was quite high; ten said it was red, three called it dark red and one said it was purple. For the second vehicle, denoted ‘Car 2’, three called it dark blue, three called blue, and two called it blue-grey. Colours blue-black, navy, black, grey, purple and blue-silver were used as the descriptive colour from each of the remaining observers. Even though there is a strong indication that the car is blue, it would be difficult to determine the colour of the car based on the free text descriptions. Also in this experiment, we are questioning a comparatively large number of observers. In most police incident there are very few witnesses, and if one said the vehicle they saw was blue, the other said black, the other purple, it would be difficult to know which colour of car to should be looked for.

Subsequently, the observers were asked then to choose the colour of the car from a list of colour names. For Car 1, seven chose dark red, four chose red; burgundy, magenta and light brown were each chosen once. For Car 2, dark blue was chosen seven times, navy was chosen five; black and grey were each chosen once. It was noted that on several occasions, people chose some colour in the free text descriptions and chose totally different colours from the lists.

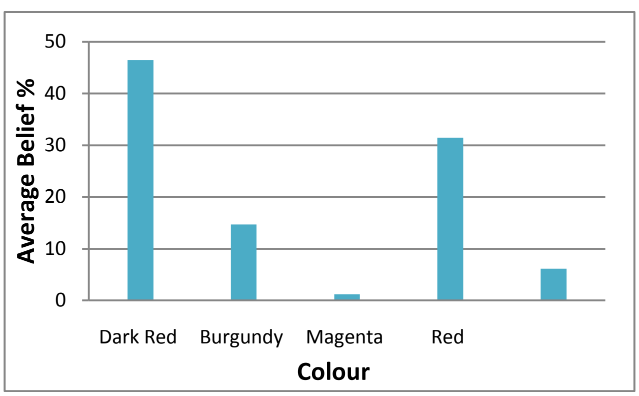

Figure 1 Observers’ percentage mean answers for Car1.

The third form of question asked observers to indicate the observers were asked to indicate more than one answer and put the percentage of their belief in each colour, effectively allowing for mixtures. The average percentage of belief given to each colour is plotted in Figure 1 for Car 1 and Figure 2 for Car 2. It is evident that the observers’ opinions are divided among the colours. The discrete TBM can be used to fuse the conflicting opinion the results are shown in Figure 3 for Car 1 and in blue in Figure 4 for Car 2.

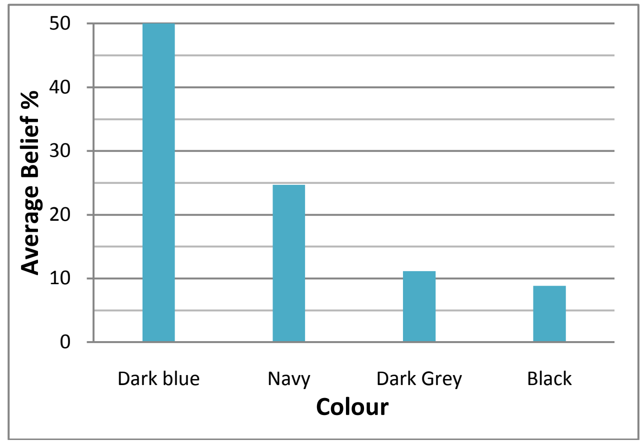

Figure 2 Observers’ percentage mean answers for Car2.

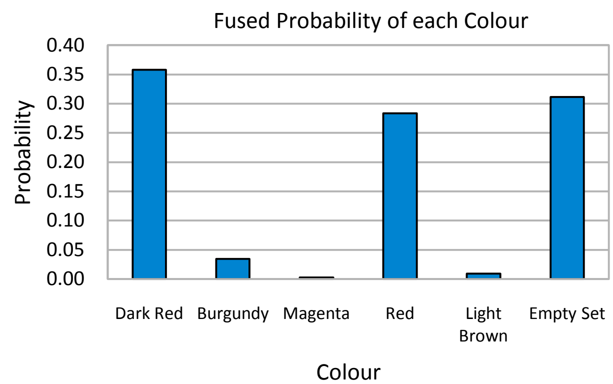

Figure 3 BetP from the fused bba of all observers for Car1.

In both cases the fusion mechanism classified correctly the colour of the car, as shown by the max *BetP* shown as the probability in the graph. The probability in the empty set indicates that there is a disagreement in the observers’ beliefs. For Car 1 this also highlights the need for an ontology, to understand how the descriptive words used to express the colours are related. This would indicate that the two colours are essentially the same, as *red* is the super-class of *dark red*, and consequently the two assigned probabilities support the same hypothesis.

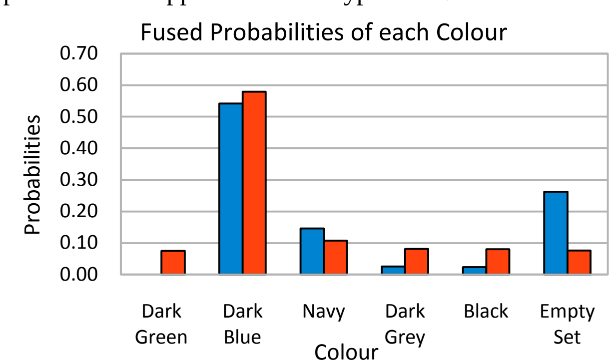

Figure 4 BetP from the fused bba (blue), and the fused bba and bbd (red) of all observers for Car2.

The mean discrepancy for the observer opinions of the real colour of the vehicle is 54% for Car 1 (with standard deviation of 31), and 46% for Car 2 (with standard deviation of 29). It can be seen, that by fusing together the different formats of the same answer, random errors can be eliminated, as is shown in Figure 4. The blue bars are the BetP from the fused bba (blue), and the red bars is the fused bba and bbd of all observers for Car2. The fusion of the individual observer opinions is performed using the discrete TBM. As can be seen there is a strong indication that the car is dark blue, but the conflict, as indicated by the empty set, is quite large. The next step is to fuse, in the continuous domain, the palette colours (their RGB value) for all the observers. Then the *bba*s are fused together with the *bbd*s, and the result is shown in Figure 4 in red. One can see that the conflict is now resolved and the empty set value is very low. There is therefore a clear indication that it is dark blue.

6.2 Human Height Estimation

In the second part of the questionnaire the 14 observers were asked to estimate the height of two subjects, Person 1 and Person 2, and classify their height in one of the categories (very short, etc). The experiment took place twice to test intra-observer variability, and the results are described in detail below.

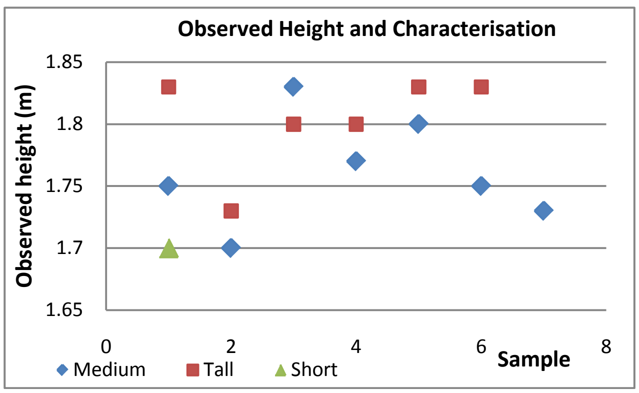

Figure 5 Characterisation of Person 2 vs. estimated height.

During the first observation (A), twelve people described Person 1 as *medium height*, while one called him *short* and one *tall*. During the second observation (B) only eight called him *medium height*, while five called him *short* and one *tall.* During the first observation (A), one observer called Person 2 *short*, while eight called him *medium height* and five *tall*. During the second observation (B) again one observer called him *short*, while seven *medium height*, and six *tall*. As shown in Figure 5, the variability among the responses is quite high the height 1.7m is simultaneously described as '*short* and *medium height*', while 1.73m is described as *tall*. A clearer distinction between the estimates of height and the height category assigned would be expected. Also a person gave Person 1 and Person 2 exactly the same height, but called one *medium height* and the other *tall*! This highlights how variable such a simple task is in a controlled environment, one can imagine how the data will look in the real world, when witnesses are asked to describe an event that took place six months ago in the past!

In order to determine how people classify humans according to their height, a height range for each class was taken. So when a group of people classified the same person at the same height class, all the estimated height values from this group were used to give the range of that class. The results are shown in Table 1. The mean height for the class *Medium* from the P1 Obs A is $M = 1.74\,metres$ (5 feet 8.5 inches) and the standard deviation $STD = 0.04$, while from the second person the same group of people gave the mean for *Medium* at $M = 1.79\,metres$. Note that the mean for the *Tall* class $M = 1.78\,metres$. This means that the people that classified Person 2 as *tall*, estimated lower height for him than those that called him *medium height*. On their second attempt Obs B, the same group changed considerably their estimations reducing the mean for *Medium* at $M = 1.77\,metres$, while increasing the mean for the *Tall* class $M = 1.8\,metres$ as shown in Table 1.

Table 1 Height for different classes from Observations.

	P1 Obs A		**P1 Obs B**		**P2 Obs A**		**P2 Obs B**	
	M	***STD***	***M***	***STD***	***M***	***STD***	***M***	***STD***
Short			1.73	0.03				
Medium	1.74	0.04	1.74	0.03	1.79	0.04	1.77	0.05
Tall					1.78	0.04	1.8	0.04

From Table 1 it is evident that there is great variability between the repeated observations. The precision of the answers is then calculated, and is used as a measure of inter-observer variability, giving how closely the answers of the individual observers match with each other. For person 1, the mean estimated height is 1.73 m (with standard deviation 0.04), indicating a low precision, since standard deviation of 4 cm is high. During the second observation the mean estimated height is 1.73 m (with standard deviation 0.03), giving improvement in precision. For person 2, the mean estimated height is 1.78 m (with standard deviation 0.04), indicating a low precision. During the second observation the mean estimated height is 1.78 m (with standard deviation 0.05), decreasing precision.

As a measure of accuracy the difference of the estimated value with the actual value is calculated. The mean difference of the real height of Person 1 to the observers' estimates is -1.2 cms (with standard deviation 3.612). During the second observation, the mean difference became -1.15 cms (with standard deviation 3). For Person 2, the mean difference of the real height to the observers' estimates is -1.07 cms (with standard deviation 4.23). During the second observation, the mean difference improve to -0.85 cms (but with standard deviation of 4.86).

The above analysis showed that the human observations were not able to accurately characterise the person's height. An attempt was made to detect the correlation between observer's height and error in observation, but the results were inconclusive. However, in case the observer was quite short , they would rate both men *tall*, while if the observer was quite tall, they would rate them as *short*. This was expected, the error difference however, did not seem to correlate with the observers height.

In the specific experiment, the observers were asked to rate the height of two people they knew quite well. The error if they only saw the person once and tried to recall from memory their height is expected to be much higher. So if this information was to be used to query a database system for people with similar height from information from witness statements, the process would not be expected to bring any meaningful results. There is therefore the need to process and standardise this information further before it becomes useable.

At the next step, the observers were asked to give the average height for each of the five categories. Again, in some cases, the range they gave did not correspond to the class and the height they estimated for the two men. There was however, greater correlation between the ranges that the observers gave. The mean and standard deviation of their answers determined the density function to be used for the TBM, to classify each person in one of the categories.

Table 2 Range of classes' height from Observers.

	Observer's Range	
	Mean	***STD***
Very Short	1.57	0.07
Short	1.65	0.06
Medium	1.73	0.05
Tall	1.84	0.04
Very Tall	1.93	0.04

The continuous TBM (cTBM) was used to classify the two men's heights based on the densities created above. The results give a clear indication of the most probable class that the two men belong to.

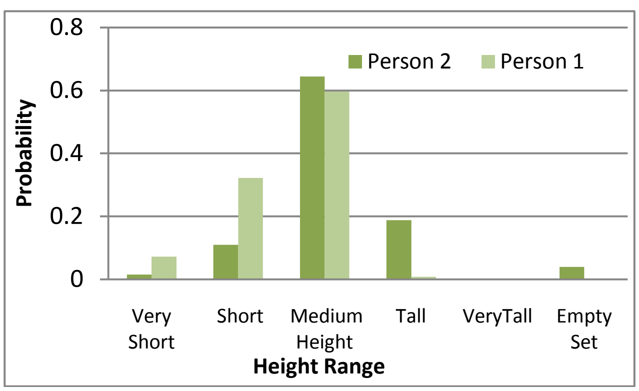

Figure 6 Classification of a Persons height

As shown in Figure 6, both men according to their height, fall into the medium height category. Person 1 is in the lower range with 1.72 m, while Person 2 at high end of the range with 1.77m. The fused information can be used to indicate more accurately to the intelligence analyst the height of the person in question.

7 Conclusion and Future Work

Judgement biases in humans even in simple tasks are significant, as shown in this case study. Extensive research with a much greater sample of observers is required. The aim of this paper was to demonstrate the complexity of including human observations as evidence and how the TBM and cTBM can be used to fuse the inconsistent information and minimise the observation errors.

A short case study was used to investigate various aspects of observers' testimony and their perceptions of colour and human height. Large inter- observer and intra-observer variabilities were noted among the results. Fusion was used to combine the different observation and combine the disagreeing opinions. The fusion result has proved to perform better than simple combinations of inconsistent information. Smaller accuracy errors were also noted. Also it demonstrated that multiple formats of answers are required to capture more accurate information from observers and eliminate random errors. Free text should not be used, the answers need to be formalised and the observers to choose their answers. Multiple methods to get various format descriptions should be used.

Future work would include the fusion of different attributes for the same entity. For example, the person's height will be fused with the colour of their shirt, etc.

References

[1] Stampouli, D.; Vincen, D.; Powell, G., *Situation Assessment for a Centralised Intelligence Fusion Framework for Emergency Services*. 6. 12th Conference on Information Fusion, Seattle US, 2009.

[2] Hall, D.L.; McNeese, M.; Llinas, J.; Mullen, T. *A framework for dynamic hard/soft fusion*, Information Fusion, 2008 11th International Conference on , Koln, 2008

[3] Pravia, M.A.; Babko-Malaya, O.; Schneider, M.K.; White, J.V.; Chong, C.-Y.; Willsky, A.S.; , *Lessons learned in the creation of a data set for hard/soft information fusion*, 12th Conference on Information Fusion, Seattle US, 2009

[4] Wright, E.J.; Laskey, K.B.; *Credibility Models for Multi-Source Fusion*, 9th International Conference on Information Fusion , 2006

[5] National Intelligence Model - Cover, Contents & Introduction, http://police.homeoffice.gov.uk/publications/operational-policing/nim-introduction.html

[6] P. Smets. *Belief functions: the disjunctive rule of combination and the generalised Bayesian theorem*, International Journal of Approximate Reasoning, 9:1–35, 1993.

[7] P. Smets. *Belief functions on real numbers*. International journal of approximate reasoning, 40(3):181–223, 2005.

[8] G. Shafer. *A mathematic theory of evidence*. Princeton University Press, 1976.

[9] P. Smets. *The transferable belief model for quantified belief representation*. In P. Smets, editor, Quantified Representation of Uncertainty & Imprecision, volume 1 of Handbook of Defeasible Reasoning and Uncertainty Management Systems, pages 267–301. Kluwer, Doordrecht, 1998.

[10] F. Caron, B. Ristic, E. Duflos, and P. Vanheeghe. *Least committed basic belief density induced by a multivariate Gaussian: Formulation with applications.* International Journal of Approximate Reasoning Volume 48, Issue 2, Pages 419-436, June 2008.

[11] C. A. Meissner and S M. Kassin. *He's guilty!: Investigator bias in judgments of truth and deception* Law & Human Behavior (2002).

[12] C. A. Meissner, J. R. Evans, Susan E. Brandon, Melissa B. Russano, and Steven M. Kleinman. *Criminal versus HUMINT interrogations: The importance of psychological science to improving interrogative practice*. Journal of Psychiatry & Law (2010).

[13] Stevens, Stanley Smith *On the theory of scales of measurement*. Science 103: 677–68 (1946).

[14] Powell, G.; Marshall, D.; Smets, P.; Ristic, B.; Maskell, S. *Joint Tracking and Classification of Airbourne Objects using Particle Filters and the Continuous Transferable Belief Model*. International Conference on Information Fusion, (2006)