Truth Discovery under Resource Constraints

Anthony Anietie Etuk

A dissertation submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
of the
University of Aberdeen.

Department of Computing Science

2015
Declaration

No portion of the work contained in this document has been submitted in support of an application for a degree or qualification of this or any other university or other institution of learning. All verbatim extracts have been distinguished by quotation marks, and all sources of information have been specifically acknowledged.

Signed:

Date: March 27, 2015
Abstract

Social computing initiatives that mark a shift from personal computing towards computations involving collective action, are driving a dramatic evolution in modern decision-making. Decision-makers or stakeholders can now tap into the power of tremendous numbers and varieties of information sources (crowds), capable of providing information for decisions that could impact individual or collective well-being. More information sources does not necessarily translate to better information quality, however. Social influence in online environments, for example, may bias collective opinions. In addition, querying information sources may be costly, in terms of energy, bandwidth, delay overheads, etc., in real-world applications.

In this research, we propose a general approach for truth discovery in resource constrained environments, where there is uncertainty regarding the trustworthiness of sources. First, we present a model of diversity, which allows a decision-maker to form groups, made up of sources likely to provide similar reports. We demonstrate that this mechanism is able to identify different forms of dependencies among information sources, and hence has the potential to mitigate the risk of double-counting evidence due to correlated biases among information sources.

Secondly, we present a sampling decision-making model, which combines source diversification and reinforcement learning to drive sampling strategy. We demonstrate that this mechanism is effective in guiding sampling decisions given different task constraints or information needs. We evaluate our model by comparing it with algorithms representing classes of existing approaches reported in the literature.
Acknowledgements

I am grateful to God Almighty, for His love and mercy in my life.

My family has been a pillar of love and support. Thank you Rita, Emediong, and Utibe Abasi. Also, many thanks to my mum, siblings, in-laws, and other relatives, for their prayers and encouragement.

I would like to thank my supervisors, Tim Norman and Murat Şensoy. I am deeply grateful for your guidance, advice, support, and above all, kindness. Also, I am grateful to Nir Oren for his invaluable contributions to this research and for always keeping an open door. I’d also like to thank my collaborators and friends at IBM T. J. Watson Research, Chatschik Bisdikian (of blessed memory) and Mudhakar Srivatsa, for their contributions to this research, and for their immense support during my time at IBM.

This research has been made less stressful, thanks to the strong network of friends I have been privileged to have, both within the university and beyond. I’d like to thank all the lovely people I came to associate with in dot.rural and Meston. In particular, I would like to thank members of 917: Peter, Andy, Gina, Mukta, Ramona, Fiona, Ruth, Danny, Danilo, Hien, Akanimo, Lizzy, and Claire, and 245: Federico, Gurleen, Luca, Hengfei, Gideon, Andrew, and Michael. You have all inspired me in a variety of ways. Special thanks to Alice Toniolo for her guidance and friendship from the very beginning of this research to the end. Finally, I would like to thank Gloria and John for their encouragement; Gertrude and Dick, and big brother Phil, for finding time to proofread parts of this thesis; Aberdeen African Choir family, for being ever present with love and cheer.

This research was sponsored by the Petroleum Technology Development Fund (PTDF) Overseas Scholarship Scheme (OSS) Nigeria, and the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorised to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.
# Contents

## 1 Introduction
1.1 Problem Context .................................................. 16
1.2 Motivation ......................................................... 17
1.3 Contributions ...................................................... 19
1.4 Thesis Outline ..................................................... 20

## 2 Related Work
2.1 Multi-Source Integration .......................................... 22
  2.1.1 What is Multi-Source Integration? ........................... 23
  2.1.2 Instances of Multi-Source Integration ................. 24
2.2 Information Fusion ................................................ 26
  2.2.1 General Fusion Methods ....................................... 27
  2.2.2 Dealing with Biases ........................................... 29
2.3 Trust and Reputation .............................................. 31
2.4 Source Selection .................................................. 34
  2.4.1 Crowd-Based Approaches .................................... 35
  2.4.2 Sampling-Based Approaches ................................. 36
  2.4.3 Decision-Theoretic Approaches ............................. 38
  2.4.4 Dealing with Dependencies .................................. 41
  2.4.5 Diversity ....................................................... 42
2.5 Summary .......................................................... 42

## 3 Background
3.1 Subjective Logic .................................................. 44
  3.1.1 Dempster-Shafer Theory ...................................... 45
  3.1.2 Opinion Representation ...................................... 46
  3.1.3 The Beta Distribution ....................................... 48
  3.1.4 Evidence Aggregation ....................................... 50
  3.1.5 SL Operators .................................................. 52
3.2 Decision Tree Learning .......................................... 53
  3.2.1 Decision Tree Representation ............................... 54
  3.2.2 Constructing Decision Trees ................................. 54
  3.2.3 Model Tree Learning ......................................... 56
3.3 Cluster Analysis .................................................. 58
## CONTENTS

3.3.1 Components of Clustering ........................................... 58
3.3.2 Pattern Representation ............................................. 59
3.3.3 Similarity Measures ................................................. 60
3.3.4 Clustering Algorithms ............................................... 61
3.4 Reinforcement Learning ................................................ 63
3.4.1 Bandit Problems ..................................................... 64
3.4.2 Elements of Reinforcement Learning ............................... 65
3.4.3 Reinforcement Learning Methods ................................ 67
3.4.4 Action Selection Strategy .......................................... 71
3.5 Summary ................................................................. 72

4 Source Diversification .................................................... 73
4.1 The TIDY Framework ..................................................... 75
4.1.1 Source Agreement ..................................................... 77
4.1.2 Trust Assessment ..................................................... 78
4.1.3 Sampling ............................................................. 78
4.1.4 Fusion ............................................................... 79
4.2 A Realisation: TIDY₀ .................................................... 79
4.2.1 Task ................................................................. 80
4.2.2 Information Source .................................................. 80
4.2.3 Report ............................................................... 80
4.2.4 Feature ............................................................. 80
4.2.5 Computing Source Agreement and Trust .......................... 81
4.2.6 Learning a Similarity Metric ..................................... 82
4.2.7 Creating a Diversity Structure ................................... 85
4.2.8 Model Validity ....................................................... 86
4.2.9 Sampling ............................................................. 87
4.2.10 Fusion .............................................................. 87
4.3 Evaluation ............................................................... 88
4.3.1 Experimental Environment ....................................... 89
4.3.2 Results .............................................................. 92
4.4 Discussion .............................................................. 99
4.5 Summary ............................................................... 99

5 Sampling Decision-Making ............................................... 101
5.1 The DRIL Model .......................................................... 101
5.1.1 Sampling State ....................................................... 102
5.1.2 Sampling Strategy .................................................... 103
5.1.3 Reward ............................................................... 104
5.1.4 Task Constraint ....................................................... 104
5.2 A Realisation: TIDY₁ .................................................... 105
5.2.1 Learning Sampling Strategies ................................... 105
5.2.2 State Space Approximation ...................................... 108
List of Tables

3.1 Sample labels for credit rating decision tree ................. 54
4.1 Training examples ........................................... 83
4.2 Experimental parameters .................................... 89
4.3 Source profiles ................................................. 90
5.1 Utility table for sampling example ......................... 107
5.2 Experimental parameters ................................. 111
# List of Figures

1.1 A layered architecture for truth discovery ........................................ 20
2.1 An example multi-source application .................................................. 23
3.1 SL opinion triangle with example opinion .......................................... 48
3.2 Uniform PDF Beta$(p \mid 1,1)$ ............................................................. 49
3.3 Symmetric PDF Beta$(p \mid 5,5)$ ......................................................... 50
3.4 Symmetric PDF Beta$(p \mid 70,70)$ ....................................................... 50
3.5 Example skewed PDF Beta$(p \mid 40,10)$ ............................................. 51
3.6 Example decision tree for credit rating ............................................... 53
3.7 Node purity for the Income attribute ............................................... 55
3.8 Node purity for the Credit history attribute ........................................ 55
3.9 Example model tree for CPU performance ........................................ 57
3.10 Linear models for CPU model tree ................................................. 58
3.11 Data clustering .................................................................................... 59
3.12 Stages in clustering ............................................................................ 59
3.13 A taxonomy of clustering approaches ............................................... 61
3.14 An example dendrogram for a 7-object clustering ................................ 63
3.15 Basic reinforcement learning scenario. ............................................ 65
3.16 Policy evaluation and improvement sequence .................................... 69
4.1 The TIDY framework ........................................................................... 76
4.2 Feature-subgroup relationship ............................................................. 82
4.3 An induced similarity metric ............................................................... 84
4.4 Diversity structure graph for a 4-agent population ............................... 85
4.5 Diversity-based fusion ....................................................................... 88
4.6 Increasing proportion of malicious sources with different budget ($\Phi$) constraints ................................................................. 91
4.7 Increasing degree of source dependence with different budget ($\Phi$) constraints ................................................................. 95
4.8 Comparing different $\psi$ (diversity threshold) parameter values: 0.4 and 0.6 ........................................................................ 98
5.1 The DRIL model .................................................................................. 102
5.2 Reward relation in sampling example ................................................ 107
5.3 Reward and mean absolute error for $\lambda = 0.1$ .................................. 113
5.4 Reward and mean absolute error for $\lambda = 0.2$ .................................. 113
5.5 Reward and mean absolute error for $\lambda = 0.3$ .................................. 115
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.6</td>
<td>Reward and mean absolute error for $\lambda = 0.4$</td>
<td>115</td>
</tr>
<tr>
<td>5.7</td>
<td>Reward and mean absolute error for $\lambda = 0.5$</td>
<td>117</td>
</tr>
<tr>
<td>5.8</td>
<td>Reward and mean absolute error for $\lambda = 0.6$</td>
<td>117</td>
</tr>
<tr>
<td>5.9</td>
<td>Reward and mean absolute error for $\lambda = 0.7$</td>
<td>118</td>
</tr>
<tr>
<td>5.10</td>
<td>Reward and mean absolute error for $\lambda = 0.8$</td>
<td>119</td>
</tr>
<tr>
<td>5.11</td>
<td>Reward and mean absolute error for $\lambda = 0.9$</td>
<td>119</td>
</tr>
<tr>
<td>6.1</td>
<td>An illustration of a crowdsourcing process</td>
<td>126</td>
</tr>
<tr>
<td>6.2</td>
<td>Pipeline monitoring using sensor networks</td>
<td>129</td>
</tr>
</tbody>
</table>
# List of Algorithms

1. Value iteration algorithm for model-based learning. .......................... 68
2. Policy iteration algorithm for model-based learning. .......................... 68
3. A Monte Carlo algorithm based on *policy iteration*. .......................... 70
4. Q-learning Temporal Difference algorithm. ....................................... 70
5. SARSA Temporal Difference algorithm. ............................................ 71
### List of Symbols/Abbreviations

- \( \omega_x \): Opinion about proposition \( x \)
- \( b_x \): Belief about proposition \( x \)
- \( d_x \): Disbelief about proposition \( x \)
- \( u_x \): Uncertainty about proposition \( x \)
- \( a_x \): Base rate of proposition \( x \)
- \( \alpha \): Alpha parameter of a beta probability distribution function
- \( \beta \): Beta parameter of a beta probability distribution function
- \( \text{sim} \): Similarity metric
- \( S \): Finite set of states
- \( A \): Finite set of actions
- \( P_s \): State transition probability
- \( P_r \): Reward probability
- \( \pi \): Policy
- \( R_t \): Total reward at time \( t \)
- \( \gamma \): Discount factor
- \( \eta \): Learning rate
- \( \bar{r} \): Vector of rewards
- \( T_{mp} \): Temperature parameter for action selection
- \( \theta^i \): Environmental state at time \( t \)
- \( \hat{\theta}^i \): Estimate of environmental state at time \( t \)
- \( \Theta \): Set of possible values of \( \theta^i \)
- \( N \): Set of all sources
- \( V_x \): Vector containing values for source \( x \)'s features
- \( o_x \): Report received from source \( x \)
- \( O_x \): Set of all reports from source \( x \)
\( \delta \)  Confidence measure of a report
\( F \)  Set of features
\( D S \)  Diversity structure
\( G \)  Group of sources
\( v_{agr} \)  Function for assessing report agreement
\( \Pi_{agr} \)  Agreement representation model
\( \sigma \)  Source agreement
\( v_{tru} \)  Function for assessing report reliability
\( \Pi_{tru} \)  Reliability representation model
\( \tau \)  Source trustworthiness
\( L \)  Learning interval
\( \Phi \)  Budget
\( P_l \)  Population change probability
\( \psi \)  Diversity threshold
\( \delta_{agr} \)  Report agreement threshold
\( P_c \)  Degree of conformity
\( P_r \)  Probability of reliability
\( \delta_{tru} \)  Report reliability threshold
\( \lambda \)  Task constraint coefficient
\( r_{qual} \)  Quality reward
\( r_{cost} \)  Cost reward
\( SQ \)  Scalarised \( Q \)-values
\( T \)  Vector containing trust levels of groups
\( \Sigma \)  Vector containing agreement levels of groups
\( SL \)  Subjective Logic
\( DST \)  Dempster-Shafer Theory
\( DTL \)  Decision Tree Learning
\( ER \)  Expected Error Reduction
\( LM \)  Linear Model
\( NRMSE \)  Normalised Root Mean Square Error
\( MAB \)  Multi-Armed Bandit
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RL</td>
<td>Reinforcement Learning</td>
</tr>
<tr>
<td>NRMSE</td>
<td>Normalised Root Mean Square Error</td>
</tr>
<tr>
<td>MDP</td>
<td>Markov Decision Process</td>
</tr>
<tr>
<td>MC</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>TD</td>
<td>Temporal-Difference</td>
</tr>
<tr>
<td>TIDY</td>
<td>Trust in Information through Diversity</td>
</tr>
<tr>
<td>DBS</td>
<td>Diversity-Based Sampling</td>
</tr>
<tr>
<td>OBS</td>
<td>Observation-Based Sampling</td>
</tr>
<tr>
<td>MBS</td>
<td>Majority-Based Sampling</td>
</tr>
<tr>
<td>RBS</td>
<td>Random-Based Sampling</td>
</tr>
<tr>
<td>MORL</td>
<td>Multi-Criteria Reinforcement Learning</td>
</tr>
<tr>
<td>ADSS</td>
<td>Adaptive Stratified Sampling</td>
</tr>
<tr>
<td>CTRS</td>
<td>Classical Trust and Reputation Sampling</td>
</tr>
<tr>
<td>DRIL</td>
<td>Diversity modelling and Reinforcement Learning</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

In recent years, social computing has emerged as an important multi-disciplinary research theme, driven by the wealth of easily available information, and the success of modern networked systems that operate across organisational and social boundaries, facilitating interaction between entities, both human and computational. Social computing initiatives largely reflect a shift from individual computing towards computations involving collective action and social interaction, with a potential for rich exchange of information. The rationale behind this concept stems from the observation that humans are profoundly social. Social tendencies are not exclusive to human societies, however. In artificial societies, such as multi-agent systems (MAS), entities or agents within these systems must necessarily interact with one another in the pursuit of their goals (Jennings, 2000). The focus of social computing is to design systems that support useful functionality that make socially produced information available to their users (Schuler, 1994). Consequently, entities have been empowered, through a variety of enabling technologies (e.g., smartphones, the Internet), to manifest their creativity, engage in social interaction, contribute their expertise, disseminate information, and assimilate collective wisdom in ways that have a profound impact on their environments (Parameswaran and Whinston, 2007).

Being “social” implies that agents are often sensitive to the actions and behaviour of those around them, and may arrive at decisions that are shaped by their social context. The underlying drive here is that individual limitations make most agents ill-equipped to handle most decision-making tasks. As a result, agents increasingly rely on the opinions of others, especially in situations where their current knowledge is inadequate or insufficient (Belkin, 2000). For instance, a “social approach” known as crowdsourcing (Howe, 2006) that taps into the potential of a large and open crowd of agents, is increasingly being embraced as a facilitator for decision-making. An example application within this context is the use of social media, such as Twitter1, to gather relevant information about an event of interest. This shift in computing style, however, raises lots of social and technical questions, such as the criteria for measuring community or group success, the relationships between social groups and individuals, the veracity of information, the demands made on communication infrastructure, and so on.

In this research, our principal concern is to advance the frontiers in social computing, by addressing the growing requirement for techniques that can assist agents deal with the complex challenges associated with social approaches to decision-making. In particular, our focus is to enable entities to effectively support decisions about what to do by supporting the problem of

1https://www.twitter.com
1.1 Problem Context

Modern society has witnessed a tremendous surge in the number and variety of information sources capable of monitoring and disseminating information about our environment. Information sources can be soft (e.g., humans) or hard (e.g., wireless sensors), they can be structured (e.g., databases) or unstructured (e.g., open source data on the Internet). The volume, velocity, and variety of data/information from these sources, coupled with the ability to share these observations, have made it possible for decision-makers or stakeholders to improve the quality and timeliness of situational assessments. With more information sources, for instance, wider coverage can be achieved, and relevant and timely information provided to appropriate quarters for decisions that could impact our collective well-being. For example, access to firsthand accounts following the 2008 Sichuan earthquake in China was made available, within seconds of the event occurring, on the social networking site Twitter.2 (Earle et al., 2010) Social approaches like this, which allow agents to leverage the power of large numbers and varieties of sources, are increasingly being exploited, and, more importantly, depended upon for decision-making. This has led to a growing interest on how to better harness the immense opportunity and potential of “the collective”. For instance, the US Geological Survey (USGS)3 recently began to investigate how social networking platforms (e.g., Twitter) can augment earthquake response and the delivery of hazard information (Earle et al., 2012).

To effectively support decisions about what to do, one needs to support the problem of what to believe. Truth discovery is an essential activity in both human and artificial societies, whereby agents try to infer the true state of the world from potentially different and conflicting information (Kittur et al., 2007). Generally, when individuals are faced with situations in which their current knowledge is inadequate or insufficient, they may seek (and integrate) further information from a variety of sources in order to change their knowledge state. For instance, consumers often turn to Internet-based opinion forums before making a variety of purchase decisions (Dellarocas, 2006); an army officer normally relies on intelligence reports from various sources to decide whether or not to engage in a battle (Gerardi et al., 2009). This ability to tap into the “knowledge of the many” or the “wisdom of crowds” comes at a price, however. There is often a great deal of uncertainty regarding the veracity of information from the crowd. Veracity is an indication of the integrity of information, and the ability for a decision-maker to trust the information and be able to confidently use it to make important decisions. This is an inherent concern in many social decision-making contexts, such as delegation (Norman and Reed, 2002). It is generally difficult, for example, to decipher the intentions of individual sources in the crowd, nor can any assumptions be made regarding their reliability. The crowd is often regarded as a complex collection of information sources with wildly varying perspectives, reliabilities, and motivating intentions (Easley and Kleinberg, 2010). The fact that a decision-maker cannot make much assumptions regarding the behaviour of sources in many domains (e.g., in online environments), introduces a great deal of uncertainty into the system, and makes it challenging to determine what is true in the world.

---

2 https://twitter.com
3 http://www.usgs.gov
For example, there is growing evidence to suggest that online firms might manipulate consumer perceptions of their products by posting (or paying people to post) messages that unfairly praise those products (Dellarocas, 2006). This malicious activity is also often carried out by way of a Sybil attack (Douceur, 2002): the use of multiple (fake) identities in order to control a substantial fraction of a system. Recently, a book author was caught engaging in this sort of attack, by writing fake online reviews (using pseudonyms) lauding his books while criticising rivals.\footnote{http://www.telegraph.co.uk/culture/books/booknews/9521564/RJ-Ellory-admits-posting-fake-book-reviews-over-past-10-years.html}

The truth discovery or veracity problem has mainly been tackled using techniques in trust and reputation. Trust and reputation are significant notions in both human and artificial societies, and have been applied in many domains for making informed and reliable decisions (Jøsang and Ismail, 2002; Regan et al., 2006; Teacy et al., 2012). A notable example in this context is eBay.\footnote{http://www.ebay.com}
eBay is one of the earliest and best known online reputation systems. As an electronic commerce (e-commerce) platform, eBay provides both buyers and sellers a means to rate each other after each transaction. These ratings are then aggregated in some manner to help other agents make informed decisions about the selection of potential providers (Resnick and Zeckhauser, 2002). Trust allows a decision-maker or an information consumer to form expectations about the reliability of information sources or providers. This in turn may guide decisions in future interactions with those sources. For instance, a decision-maker may have known, from experience, that a certain group of sources is unreliable given that they constantly provide reports that do not reflect ground truth.\footnote{Ground truth in this context refers to an objective fact about the state of the world. This is as opposed to the subjectivity of ground truth often portrayed in most reputation models.} This expectation of reliability may form a basis for the agent’s decision concerning the group in a future interaction. The decision-maker may, for example, choose to discard reports provided by the group, avoid asking sources in the group for evidence, or transform their reports in some manner given any specific knowledge of their reporting patterns.

Our intention is not to replicate the significant body of work done in the areas of trust and reputation. Rather, our focus in this research is on truth discovery in resource-constrained environments, which presents difficulties for existing trust approaches. For instance, a general assumption of existing trust and reputation models is that querying sources is cost-free. Therefore, these approaches are usually ill-equipped to function in environments where resource use (i.e., querying a source) may affect the utility of a decision-maker.

\section*{1.2 Motivation}

Typical approaches to truth discovery is to rely on reports from as many sources as possible. The underlying assumption being that exploiting the “wisdom of crowds effect” (Surowiecki and Silverman, 2007) minimises the influence of biased opinions. Before continuing our discussion, let us consider the following scenario which captures a similar thought process, and serves to elaborate more on our motivation.

Assume a situation where a certain organisation, ABC receives an alert about an oil spill in some remote location, which might be due to a malfunctioning installation in the region. In the past week, however, ABC has received a couple of similar alerts which turned out to be false
alarms. ABC responded to these events, and as a consequence has incurred significant costs (e.g., by deploying repair engineers to the facility). The organisation is therefore wary of this latest alert, and is reluctant to make further commitments. There are, however, greater penalties for not acting. The organisation may be required to: pay for any clean-up exercise that may be needed; fix the impaired facility to avoid further damage and loss; face possible prosecution by the state or government; and pay compensation to affected parties, in addition to the adverse publicity and potential losses in share value. For these reasons, ABC decides to seek further evidence about the situation at the facility in order to decide the best or appropriate course of action. In the presence of several information sources, ABC is likely to have a higher confidence in evidence from a number of sources that together give consistent opinions about the situation. The question is: with limited capacity to query for evidence, how can a reliable group of information sources be selected, whose combined report will best reflect the true situation at the facility?

As reflected in the hypothetical scenario, capturing and distributing evidence can be costly in many real-world contexts. In distributed environments such as peer-to-peer networks, sensor networks, and pervasive computing applications, each participant is responsible for collecting and combining evidence from others due to lack of central authority or repository. In emergency response, for example, a decision-maker at some node in the network must make decisions in real time on the basis of high volumes of streaming information received from a variety of sources through different routes. Major constraints in these systems are bandwidth, delay overheads, and energy, motivating the need to minimise the number of messages exchanged. In addition, many sources such as GeoLytics, providers of demographic information; WDT, providers of weather information, and DigitalGlobe providers of space imagery and geospatial content, charge for the information products they provide. Even for sources that are free, integration often entails spending much resources to resolve conflicts and clean the data (Dong et al., 2012). Furthermore, there is often no guarantee that evidence obtained from different sources are based on direct, independent observations. Sources may likely provide (unverified) reports obtained from others. For instance, there is evidence to suggest that social influence can impact people’s opinions, leading them, in some cases, to revise their reports (Lorenz et al., 2011). A notable example that aligns with this assertion is in social sensing, where information shared by individuals (e.g., via a social network such as Facebook or Twitter) can be accessed by a wide audience (Campbell et al., 2008; Kwak et al., 2010), who may, in turn, report the same information later, possibly without any acknowledgement. This exposes one to the risk of double-counting evidence, introducing an extra challenge of distinguishing fact from rumour.

As pointed out, crowds are not necessarily wise. According to Surowiecki (2004), there are key criteria that characterise wise crowds: diversity of opinion, independence, decentralisation, and proper mechanisms for aggregating individual opinions. In Dong et al. (2012)’s view, “the more the better” might not always hold (in many environments or application contexts), it might well be the case that “less is more”.

---

7http://www.geolytics.com
8http://wdtinc.com
9http://www.digitalglobe.com
10https://www.facebook.com
11There is a recorded account at http://www.bbc.co.uk/news/uk-14490693, on how Twitter was used to spread false rumours during the England riots of 2011.
1.3 Contributions

As we have discussed, the integration of multiple information sources in environments that are characterised by resource constraints, and where social influence, or other forms of dependencies among information sources exist poses problems to truth discovery. These types of environments often violate common assumptions: (a) the greater the number of reports acquired, the better the estimate; (b) reports are independent; (c) acquiring reports is cost-free. Despite these challenges, decision-making processes that rely on collective intelligence are on the rise.

We argue that by exploiting diversity among information sources and through an intelligent sampling of the source population, we can acquire as accurate an estimate as possible under resource constraints.

Our key contributions are:

1. A general framework for multi-source integration under resource constraints (Chapter 4): An integrated framework, known as TIDY (Trust in Information through DiversitY) for truth discovery in resource-constrained environments. TIDY encapsulates a number of specialised functions (e.g., source selection, and information fusion) (Luo and Kay, 1989), which operate in a seamless manner to meet a decision-maker’s various information goals.

2. A mechanism for source diversification (Chapter 4): A mechanism for source diversification, whose aim is also to identify different forms of dependencies (e.g., copying) among information sources. A model of diversity enables us to generalise from similarity in sequences of reports from different sources to similarity of sources on the basis of their observable features. Through the realisation of the TIDY framework in this context, referred to as TIDY₀, we demonstrate that with an increasing proportion of dependent sources, an approach based on diversity performs significantly better than approaches that are not based on diversity in making accurate estimates of environmental states. When the assumption of dependency is relaxed, the model does not perform any worse than approaches that do not take diversity into consideration.

3. A decision model based on diversity and trust (Chapter 5): This mechanism builds on the previous model to learn effective sampling strategies under different task constraints. The sampling decision model allows a decision-maker to balance, in a principled manner, the trade-off between accuracy and costs of acquiring information. Through the realisation of the TIDY framework in this context, referred to as TIDY₁, we demonstrate that our sampling decision model, based on diversity and trust, yields higher utilities than approaches based only on diversity or trust.

These contributions can be illustrated by means of a layered architecture in Figure 1.1, showing how the different components of our model work together to support the problem of what to believe under resource constraints. There are various information sources (the crowd), that may be sampled for reports in order to estimate the value of some environmental state (truth discovery). Layer 1 represents the source diversification process, TIDY₀, where sources in the population are clustered into groups by exploiting evidence from their observable features and past reports. The
1.4 Thesis Outline

Our discussion in the remaining chapters is outlined as follows:

- In Chapter 2, we provide a survey of related work in multi-source integration. We discuss the problems that resource constraints and source dependence present with respect to truth discovery, and compare existing approaches that attempt to address these issues at different stages of integration.

- In Chapter 3, we provide a précis of the techniques that we exploit in building a concrete realisation of our framework. These techniques underpin the model presented in Chapters 4 and 5.

- In Chapter 4, we present our source diversification model, and evaluate its performance within a simulated multi-source system. We discuss the problems that can arise when correlated biases exist among information sources, and demonstrate how a decision-maker can protect itself against these biases in the truth discovery process.
In Chapter 5, we present our sampling decision-making model, which employs the model developed in Chapter 4 to make decisions on how best to sample the source population under different task constraints. We evaluate this model within our simulated environment, and show that by strategically sampling information sources, the utility derived by a decision-maker can be significantly increased.

In Chapter 6, we provide a critical discussion of our research, comparing it with similar work. We also reflect on ways to improve and assess it further, and in Chapter 7, we present our concluding remarks.
Chapter 2

Related Work

In this chapter, we discuss related research in the area of multi-source integration. We begin by providing some background on the notion of multi-source integration. Using example instances, we highlight challenges to the effective application of this concept in a variety of domains. In particular, we focus on the challenge of truth discovery when integrating evidence from multiple sources in large, open and dynamic environments, given that there is uncertainty regarding the trustworthiness of those sources, and a decision-maker has to operate under resource constraints. In so doing, we hope to motivate our model presented in Chapters 4 and 5.

2.1 Multi-Source Integration

The concept of multi-source integration has been around for a long time. Humans and animals have evolved the capability to use multiple senses for improved situation awareness (Hall and Llinas, 1997; Krueger et al., 1974; Gazit and Terkel, 2003). For example, hunting dogs use a combination of smell and sight in order to track their prey; the quality or appeal of an edible substance may be better assessed using a combination of sight, touch, smell, and taste; a medical consultant may order multiple diagnostic tests and seek additional peer opinions before performing a major surgical procedure.

Thanks to recent technologies, including sensing and networking capabilities, many people (and applications) can now tap into the collective on a far greater scale (Bonabeau, 2009; Ye et al., 2012; Sheth et al., 2008; Resnick et al., 2000). Indeed, the proliferation of sensing resources (both hard and soft) that provide huge amount of data about our environment, coupled with a variety of platforms (e.g., the Internet or mobile ad hoc networks) for sharing those observations constitute a paradigm shift in the way we perceive our world and arrive at decisions. Evidence of this is widespread:

- Environmental (e.g., flood detection, forest fire detection, air quality monitoring, monitoring river water level) (Mainwaring et al., 2002);

- Healthcare (e.g., patient monitoring) (Yilmaz et al., 2010);

- Military (e.g., target tracking and localisation, battlefield surveillance, chemical attack detection) (Brooks et al., 2003);

- Transportation (e.g., bus service tracking, traffic management) (Tubaishat et al., 2009);
2.1. Multi-Source Integration

A generally maintained view is that the integration of information from multiple sources can lead to better assessments than using a single source (Smith and Singh, 2006; Varshney, 1996). Also, as observed by Teacy et al. (2012), multi-source integration makes an information system robust to the absence or failure of any single source. Figure 2.1 illustrates an instance of multi-source sensing, applied in the context of a river water level monitoring.

2.1.1 What is Multi-Source Integration?

Multi-source integration is a multi-disciplinary theme that has received a significant amount of attention in many areas including signal processing, statistical estimation, control theory, and artificial intelligence (Punska, 1999). It is therefore unsurprising that its terminology is not unified. For instance, multi-source integration has often been confused with multi-source fusion. However, multi-source integration is a broader concept than multi-source fusion. Luo and Kay provide a clear distinction between the two:

Multi-source integration refers to the synergistic use of information from multiple sources to assist in the accomplishment of a task; and multi-source fusion deals with the combination of information from different sources into one representational format during any stage in the integration process (Luo and Kay, 1989).

Making this distinction is useful in the context of this research: it serves to separate the much broader issues involved in the integration of multiple sources of information at the system architecture and control level from more specific issues like fusion. The scope of multi-source integration spans across different functions such as source selection, system model, data transformation, as well as fusion (Punska, 1999; Luo and Kay, 1989). As we will see later, issues such as source selection can be clearly distinguished from other issues such as fusion or the transformation of observations from sources.
The advantages of multi-source integration can be described from different perspectives (Nakamura et al., 2007; Luo and Kay, 1989):

- **Complementarity**: Information from different sources represents different aspects of a broader scene, which may be combined to obtain a more complete view. Each source can be said to perceive features of the environment that may be independent of the features perceived by other sources. For example, in the event of an explosion in a monitored region, the combination of a sound-level report from an acoustic sensor, and a dust-level report from a human source could provide a much clearer insight into the actual state of the environment.

- **Redundancy**: A statistical advantage is gained by adding observations from independent sources about the same feature(s), possibly reported with different fidelity. Assuming the observations are combined in an optimal manner, redundancy can reduce overall uncertainty, thus increasing the confidence in the information. Variants of information redundancy have been applied for truth discovery in online transactions by pooling opinions from multiple sources to assess the reliability of a potential transaction partner (Jøsang and Ismail, 2002).

- **Cooperativeness**: Information from sources can be combined into new information which better represents the true state of the world. For instance, a pulsed radar can accurately determine a target’s range but has a limited ability to determine angular direction of the target. By contrast, an infrared imaging sensor can accurately determine the target’s angular direction, but not its range (Hall and Llinas, 2001). Cooperativeness is relevant in areas such as target localisation, where computation of a target’s location can be performed based on angle and distance information from different sources (Lipton et al., 1998). Whereas complementarity involves sensors that perceive different properties of the environment (e.g., sound, temperature), cooperativeness involves sensors providing a different or a partial view of the same scene.

### 2.1.2 Instances of Multi-Source Integration

Several information systems exist that enable the integration of multiple sources in a decision process. Most of these platforms have emerged as a result of advances in sensor and network technologies. In particular, these systems provide a medium for an agent to pool the opinions or perspectives of multiple sources (Durrant-Whyte and Henderson, 2008) or, in the words of Surowiecki and Silverman, ‘the wisdom of crowds’ (Surowiecki and Silverman, 2007) for decision-making. We provide an overview of some of the common ones, with a view to highlighting the benefits they offer. Not only that, we highlight some of the issues that must be addressed if these systems are to meet their full potential. One such class of systems that have generated a great deal of interest in recent years is sensor networks.

**Sensor Networks**

A sensor network (Rowaihy et al., 2007; Akyildiz et al., 2002) is a special type of ad hoc network composed of a large number of sensor nodes. These nodes are capable of taking measurements (e.g., acoustic, temperature) of their environment. Sensor nodes are usually deployed very close to a phenomenon of interest, although their deployment may be in an ad hoc manner. This allows random deployment in inaccessible terrains or disaster relief regions. Once deployed, the network
is often left unattended to perform monitoring and reporting functions (Yick et al., 2008). Each sensor node is equipped with a small processor and wireless communication antenna. In addition, each node is powered by a battery. The deployed nodes can communicate with each other, and with a data processing or fusion centre. An excellent survey of sensor networks and their applications can be found in Yick et al. (2008) and Akyildiz et al. (2007, 2002).

In order to achieve the goal of deployment, especially in inaccessible regions, a sensor network must last for long periods of time (e.g., months or years). However, this objective is often hampered by the limited (and generally irreplaceable) energy available to the nodes. Therefore, sensor network protocols must also focus on power conservation. Some of the applications of sensor networks include military, health, and environmental applications. Many of these applications involve mission-critical tasks, and are often constrained by limited time and resources, making it necessary to avoid unproductive sensor actions. The goal therefore, should be to select the subset of sensors that are the most decision-relevant (Zhang and Ji, 2010).

There is also the issue of uncertainty on a network. Sensor nodes are implemented using inexpensive hardware components that are highly unreliable, deployed unattended in hostile environments, and are severely constrained in energy and bandwidth. These nodes may sometimes misbehave, for instance, due to faults in software and hardware (Ganeriwal et al., 2008). A node may simply drop packets, for example, to save power. An attacker could compromise a number of sensors located in a physically insecure place, or some natural event could impair a group of sensors, thus impacting the reliability of the information they provide. Chen et al. (2009) and Walters et al. (2007) provide excellent surveys of the different reliability issues in sensor networks. Making meaningful assessments of environmental states using this medium will imply accounting for the uncertainty in the reports provided by individual nodes.

**Social Networks**

When a computer network connects people or organisations we have a social network. A social network is a set of actors or nodes (i.e., people, organisations or other entities) connected by a set of social ties or relationships (e.g., friendship, co-working) (Garton et al., 1997).

Social networks create unprecedented opportunities for information sharing and are increasingly being leveraged for decision-making. In social sensing, people act as sensors or sensor operators sharing information within social and special interest groups (Campbell et al., 2008). An example application is a participatory sensing campaign to report the location of an event (e.g., earthquake) (Sakaki et al., 2010). Crowdsourcing (Doan et al., 2011; Howe, 2006) is another significant example in this context. This process can be employed to solicit opinions or ideas from a large network of people in the form of an open call (Kittur et al., 2008). A classic example of a crowdsourcing application is the Amazon Mechanical Turk (MTurk), a web service that coordinates the supply and the demand of tasks that require human intelligence to complete (Paolacci et al., 2010). Ancillary services to MTurk such as TurkerNation and Turkopticon, provide a forum for workers to share and evaluate their experiences of employers. Web surveys can also be regarded as the “crowdsourcing” of responses from multiple sources online (Couper, 2000).
Web surveys make information collection available to the masses, thereby opening up a whole new realm of information gathering possibilities that would otherwise be impossible or extremely difficult to achieve. These applications are made popular by the ubiquity of network connectivity (e.g., via smartphones), and the means of information dissemination via social network platforms such as Facebook\(^4\) and Twitter (Kwak et al., 2010; Joinson, 2008).\(^5\)

One of the most basic notions governing the structure of social networks is homophily: the principle that people tend to behave similarly to their friends (Easley and Kleinberg, 2010; McPherson et al., 2001). Friendship, according to Easley and Kleinberg, can be defined along different social characteristics including racial or ethnic dimensions, age-group, occupation, interests, beliefs, etc. Homophily impacts people’s social worlds in a way that has powerful implications to the information they share, the attitudes they form, and the interactions they experience (McPherson et al., 2001). Thus, information flowing through a social network tends to be localised, setting the stage for the emergence of cliques or subgroups (Garton et al., 1997; Dunbar and Spoors, 1995). As mentioned, there are significant implications. For instance, opinions from individuals can easily be influenced by others or the (logical) subgroups to which they belong. This may lead to correlated biases in opinions from groups members (Cha et al., 2010; Watts and Dodds, 2007; DeMarzo et al., 2003).

Perhaps the best way to view a social network in terms of the integration of multiple sources of information, is in the words of Easley and Kleinberg:

“…crowded with an array of information sources of wildly varying perspectives, reliabilities, and motivating intentions.” (Easley and Kleinberg, 2010)

Therefore, understanding any one piece of information may very much depend on knowledge of how the information relates to other pieces of information on the network (Easley and Kleinberg, 2010). This view is consistent with the fact that information shared by individuals via a social network (e.g., Facebook, Twitter) can be accessed by a wide audience (Campbell et al., 2008; Kwak et al., 2010), who may, in turn, report the same information later, possibly without any acknowledgement, thus, making it difficult to distinguish fact from rumour.

As highlighted, the problem of multi-source integration is often not one of shortage of information, but rather one of how to make sense out of the diverse and possibly conflicting opinions. This generally entails fusing the different opinions or viewpoints into a single estimate that can be used for decision-making. This problem has been widely studied under information fusion.

### 2.2 Information Fusion

Information fusion, also called data fusion, deals with the combination of multiple sources to obtain improved information: cheaper, better quality or greater relevance. Hall and Llinas defines information fusion as:

“The process of combining data or information to estimate or predict entity states.”

(Hall and Llinas, 2001)

\(^4\)https://www.facebook.com
\(^5\)https://www.twitter.com
2.2. INFORMATION FUSION

The target entity may, for example, represent some aspect of the world (e.g., river water level, pressure in a pipeline). The objective might also be to estimate or predict the behaviour of an individual upon whom one’s well-being is dependent (Gambetta, 2000). Regardless of the underlying information structure, the fusion process can be expressed in a general sense:

\[ \Theta = F(x_1, x_2, \ldots, x_n), \]  

(2.1)

where \( x_i \) denotes a single report from an information source about the state of the world, and \( \Theta \) is the derived estimate upon which a decision may be based. The fusion function, \( F \), depends on the chosen fusion method. These methods are generally based on several criteria such as the purpose of fusion or type of data (Nakamura et al., 2007).

2.2.1 General Fusion Methods

Some of the common and widely used fusion methods are Bayesian inference (Box and Tiao, 2011); Dempster-Shafer inference (Shafer, 1976); Weighted average (DeGroot, 1974).

**Bayesian Inference**

Bayesian Inference (BI) provides a principled way of combining new evidence with prior beliefs according to the rules of probability theory. The unknown state or variable of interest is modeled as a random variable. Degrees of belief about the variable are represented by the *a priori* probability, conditional probability of an observation given a hypothesis, and the *a posteriori* probability. Bayes’ theorem provides a relationship between these three probabilities, and can be used to make subjective estimates of belief in the hypothesis. Fusion is usually performed by Bayes’ rule, which, under the condition of source independence, is reduced to a product of all pieces of evidence. Bayes’ rule is given as:

\[ p(\theta \mid H, S_1, \ldots, S_m) \propto p(S_1, \ldots, S_m \mid \theta, H) \ p(\theta \mid H), \]  

(2.2)

where \( \theta \) denotes the variable of interest; \( m \) denotes the number of sources; \( p(\theta \mid H) \) denotes the decision-maker’s prior probability for \( \theta \) given its knowledge \( H \); \( p(\theta \mid H, S_1, \ldots, S_m) \) denotes the decision-maker’s posterior probability for \( \theta \) given knowledge \( H \) and the sources’ opinions \( S_1, \ldots, S_m \); and \( p(S_1, \ldots, S_m \mid \theta, H) \) denotes the likelihood of the sources’ opinions given \( \theta \) and \( H \).

Decisions are usually made on the *posteriori* probabilities. Inference can be carried out iteratively: after observing some evidence. In this case, the resulting posterior probability can be treated as a prior, and a new conclusion or posterior reached given new evidence. This allows Bayesian principles to be applied to various kinds of information integration settings: whether evidence is obtained all at once or acquired over time.

An example application of BI can be found in the work of He et al. (2006), where the method is applied on a social network to study the problem of information disclosure. The authors model a social network as a Bayesian network, and use this model to infer attributes of a person (e.g., age) by combining pieces of seemingly innocuous or unrelated evidence from his/her friends’ attributes. Bayesian inference has also been widely applied in sensor networks. For instance, Krishnamachari and Iyengar (2004) adopt this method for event detection based on binary signals (0 and 1) reported by a group of sensors. To distinguish between faulty sensor measurements, the authors assume that true sensor readings are correlated, while sensor faults are likely to be
2.2. INFORMATION FUSION

uncorrelated. Thus, the work strongly relies on neighbourhood nodes (or consensus) to detect bogus reports. Fox et al. (2003) apply Bayesian-filter techniques in pervasive computing (Satyanarayanan, 2001) to estimate, for example, a person’s location. Their approach is a sequential model, which iteratively estimates (and refines) the belief over possible states (person’s or object’s location) as new evidence is integrated.

One challenge of using the Bayesian method is that it requires the specification of a prior for all unknown parameters. This can and, in fact, should be done when there is concrete prior knowledge about the parameters. For example, suppose one has an extremely small prior knowledge or belief in a hypothesis, say, existence of flying elephants. If a source provides evidence of a flying elephant, then the person’s posterior probability would nevertheless be small. In many cases, however, prior knowledge of the state of the world is either vague, or non-existent. That makes it difficult to specify a unique prior distribution. For example, different analysts or decision-makers having different background knowledge or opinions, may suggest different priors, and arrive at different answers. This issue is often overcome by using certain non-informative (or uniform) priors (Bernardo, 1979).

Still within the context of the Bayesian approach, Bayesian classifiers or naive Bayes classifiers (Rish, 2001) are simple probabilistic classifiers that assign the most likely class to a given example described by its feature. Although Bayesian classifiers have proven effective in many practical applications, including medical diagnosis, their strong (naive) independence assumptions between the features make them unsuitable in many real-world contexts.

**Dempster-Shafer Inference**

Dempster-Shafer (DS) inference, which is based on the Dempster-Shafer belief theory, is a generalisation of the Bayesian approach. Whereas the Bayesian approach focuses on assigning probabilities to each proposition of interest, DS assigns non-negative weights (belief masses) to each combination of events, and allows explicit representation of uncertainty or ignorance. For example, let’s say we are considering the proposition that the colour of a river water is one of three colours, green, brown or blue. A Bayesian approach might assign probabilities individually to each of three possibilities, green, brown and blue, say as \{0.2, 0.3, 0.5\}. DS on the other hand, would assign weights to each of the eight possibilities: \{\emptyset, \{green\}, \{brown\}, \{blue\}, \{green, brown\}, \{green, blue\}, \{brown, blue\}, \{green, brown, blue\}\}. DS does not require an exhaustive set of hypotheses to be defined (Hall and McMullen, 2004), which is useful factor when modelling situations in which it is not possible to exhaustively list all possible states of an entity. Dempster-Shafer inference allows for the representation of both imprecision and uncertainty in sources’ reports using plausibility and belief function, both of which are derived from a mass function or basic probability assignment. A more careful treatment of DS is provided in Chapter 3. However, in contrast to BI, DS allows the combination of information provided by different types of sources (Nakamura et al., 2007). Thus, the fusion rule can handle information at different levels of detail from different sources. Furthermore, DS doesn’t need the assignment of prior probabilities to unknown propositions. Instead, probability masses are assigned only when there is evidence supporting a proposition.

Bloch (1996) applies DS for medical image processing. In particular, the author demonstrates how DS can be applied to classify brain tissues by fusing images obtained from several imaging
2.2. INFORMATION FUSION

techniques. In a different application context, Basir and Yuan (2007) demonstrate how DS can be applied to engine diagnostics. Multi-sensor information such as vibration, sound, and pressure are combined in order to identify engine faults. In this problem, information obtained from different sensors are considered as different pieces of evidence for the DS fusion process. Fard and Ester (2009) apply DS in the context of social networks to perform fusion of different evidence for the discovery of criminal groups and suspicious persons.

Dempster-Shafer inference has its limitations. Zadeh (1979) points out that DS may lead to counter-intuitive results when there is a high degree of conflict. For example, if one source reports that the colour of the river water is green with 99 percent belief, while a second source reports that the colour of the river water is brown equally with 99 percent belief, but both have a 1 percent belief that it is blue. By applying DS fusion rule to combine these two sets of evidence, we obtain a combined opinion that the river water colour is blue. This result is due to the fact that the opinions (however weak) that the river water colour is blue is the only evidence that does not conflict, but the result is counterintuitive.

Weighted Average

Weighted average is perhaps one of the simplest and most intuitive fusion methods. The majority of the weighted methods are based on consensus theory (DeGroot, 1974), or linear opinion pooling (Jacobs, 1995). The estimated variable is often derived as:

\[ \Theta = \sum_{i=1}^{n} w_i P_i, \]

where \( P_i \) is the subjective probability reported by source \( i \), and \( w_i \) is a linear coefficient or weight that is used to normalise the report from the source. A necessary condition for the definition in Equation 2.3 is that the weights sum to one (Jacobs, 1995). The voting method described in Smith and Singh (2006) is a variant of the weighted average method. Here, it is assumed that the observations from the sources carry equal weight, and the hypothesis which is considered the most likely by the highest number of sources is chosen. Other instances and variants of the weighted average can be found in Ross et al. (2006) and Xiao et al. (2005).

DeGroot (1974) identifies the major issue associated with weighted average as that of deciding suitable weights for individual contributions. Benediktsson and Kanellopoulos (1999) suggest that weights should reflect the “goodness” of information sources, such that high weights are assigned to the reports from sources that contribute to high accuracy in the fusion result. The authors further suggest that some reliability measure of sources can be used as a criterion for heuristic weight selection.

A good survey on different multi-source fusion schemes, including the popular Kalman filter (Welch and Bishop, 1995) method widely used in signal processing, can be found in Smith and Singh (2006) and Nakamura et al. (2007).

2.2.2 Dealing with Biases

The majority of fusion methods are based on optimistic assumptions about the reliability of sources (Rogova and Nimier, 2004). For instance, traditional multi-source fusion applications rely on information from sources that are reasonably well known, and whose error models are understood (Wright and Laskey, 2006; Luo and Kay, 1989). Combining opinions from multiple sources in the
2.2. Information Fusion

expanded landscape of large, open and dynamic systems (e.g., the Web) must be robust to different kinds of biases. Information sources may be deceptive, inaccurate, or simply uncooperative (often due to self-interested motives). The sources may be operating individually or may work in concert (e.g., collusion) to create a more significant impact in the system. Easley and Kleinberg (2010) recognise these challenges as (a rather unfortunate) hallmark of the complex “connectedness” of modern society.

A number of approaches exist for dealing with source bias. These can generally be classified as either *endogenous* or *exogenous* (Whitby et al., 2004):

- **Endogenous**: this technique uses the statistical properties of the reports themselves as a basis for assessment. The underlying assumption is that majority of the sources are likely to be reliable. Thus, identifying unreliable reports reduces to identifying those reports that differ significantly from mainstream opinion.

- **Exogenous**: this technique focuses on learning about the reliability or trustworthiness of individual sources, generally based on the accuracy of their reports. Reports are then weighted according to the trust or perceived level of accuracy of the respective sources. In this context, it is assumed that sources with low reputation are likely to provide inaccurate reports and vice versa.

Endogenous techniques involve statistically analysing the distribution of reports provided by the sources, and evaluating each one according to its deviation from the consensus. Outlying reports, by this metric, are considered inaccurate or biased and thus filtered out. For example, Dellarocas (2000) uses a cluster filtering approach, in a binary report model, to separate reports that are unduly (or unfairly) optimistic about the truth of a proposition. However, this approach is unable to handle unfairly pessimistic reports. Whitby et al. (2004) propose another method that operates in the binary space, and uses an iterative filtering technique based on the Beta distribution to discard reports that do not comply with the reports of the majority. iCLUB (Liu et al., 2011) is also based on clustering, but, unlike the previous two approaches, iCLUB supports multi-nominal reports, which is a richer representation model to the binomial model. Yin et al. (2008) proposed TruthFinder, which utilises an iterative technique to identify the true situation in the world. Their approach is based on the assumption that an information source is reliable if it often provides true information, and a piece of information is likely to be reliable or true if it is provided by many trustworthy sources. In other words, TruthFinder also assumes that information provided by the majority of the sources is reliable. The approach proposed by Wang et al. (2012) is similar to TruthFinder in that the corroboration in sources’ reports is used to discover the true value, and this knowledge is then used to determine the reliability of the sources. The authors make use of the Expectation-Maximisation algorithm (Flach, 2012, p. 288). There are also a number of general techniques for outlier detection with strong statistical background (Chandola et al., 2009), which could be applied for dealing with biases in more general cases.

One advantage of the endogenous technique is that it does not require any additional information, such as extensive knowledge of the sources, and so is appropriate in dynamic settings when the source population changes rapidly. However, by making a strong assumption about the
2.3. Trust and Reputation

Exogenous approaches are mostly variants of trust and reputation systems. In particular, these approaches build trust models of information sources, and use these models as a basis for discounting or normalising their reports. In the next section, we highlight some relevant trust and reputation mechanisms, discussing their approaches for dealing with biases. An extensive discussion on the subject, as well as a survey of trust and reputation models can be found in Ramchurn et al. (2004) and Sabater and Sierra (2005). Also, Jøsang et al. (2007) provides an excellent survey on the application of trust and reputation for online service selection.

2.3 Trust and Reputation

Trust and reputation are at the centre of most information systems and decision-making processes. Marsh did seminal work, in his PhD thesis, on the formalisation of the concept of trust in artificial societies (Marsh, 1994). Ever since, many computational trust models have been proposed. The focus of many such models have been to address the problem of uncertainty and to improve the quality of service (Şensoy et al., 2009; Jøsang et al., 2007). In the context of truth discovery, using evidence from multiple sources, trust and reputation may serve as relevant tools for the “transformation” of information from these sources before being fused.

A universal consensus about the notion of trust is lacking, perhaps, due to its multi-disciplinary dimensions (McKnight and Chervany, 1996). In the context of this research, we adopt a view similar to that of Nevell et al. (2010): a measure of the degree of belief in the capability of an information source to provide information that conforms to fact. This may be considered as a specialisation of one of the more general and widely used definitions of trust in both human and artificial systems provided by Gambetta:

“...a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both before he can monitor such action (or independently of his capacity ever to be able to monitor it) and in a context in which it affects his own action.” (Gambetta, 2000)

A large number of trust models adopt a probabilistic view of trust. According to Jaynes (2003), probability theory provides a complete set of rules for reasoning about degrees of uncertainty that guarantees logically consistent inference about a subject. One of the earliest examples of a probabilistic trust model is the Beta Reputation System (BRS) (Jøsang and Ismail, 2002). BRS estimates the likelihood of a binary proposition such as “source i is trustworthy”, using Beta probability density functions. In particular, BRS represents a source’s ability to provide accurate reports as a binary random variable. That is, either the source is reliable or unreliable. A source’s perceived reliability or trustworthiness is then modelled by the probability that it will provide accurate reports when queried in the future. For example, if a source, i, is queried ten times for information, resulting in a 0.3 probability of reliability, then, on average, it is expected that i would
provide reliable reports on three of those occasions, and unreliable reports on seven. This probability can be estimated using evidence from (past) interactions with the source. In some settings, it may be possible and even necessary to pool opinions (i.e., reputation) from other sources in order to enrich the assessment of the trustworthiness of a source for which a decision-maker has insufficient direct experience.

Although the approach adopted by BRS to handle biases is to filter out reports that deviate from mainstream opinion (Whitby et al., 2004), the computed probability value can be used as a weighting factor for reports (Sensoy et al., 2013). A similar step is taken in Subjective Logic (Jøsang, 2013), where a decision-maker represents opinions about the trustworthiness of an information source as a triple: belief, disbelief, and uncertainty. This metric can then be used to discount the source’s opinion about the state of the world (Zhang et al., 2006).

TRAVOS (Teacy et al., 2006) is another probabilistic approach, and is similar to BRS in that it uses a Beta distribution to model the trustworthiness of a source. Though based on probability, TRAVOS uses heuristics to normalise reports from third-party information sources. In particular, the approach assumes that reports that do not align with the personal observations of the decision-maker are biased, thereby reducing their value or effect before fusion. This assumption can be problematic. For instance, in line with Gambetta’s definition of trust (presented above), it can be assumed that a decision-maker may not always be in a position or have the capacity to monitor the state of the world for which its interest lies. Even if this were the case, why would an agent request opinions of others if it has sufficient capability to monitor its environment? An agent’s view of the world being thus imperfect or subjective, can it then be regarded as a good enough yardstick for assessing the reliability of other sources with respect to reporting an objective fact?

Yu and Singh (2002) propose an approach based on Dempster-Shafer theory. An interaction with a source is represented by a set of values that reflects the quality of that interaction. They represent their model of a source using three quality of service (QoS) categories: trustworthy, uncertain, and untrustworthy. Two thresholds (upper and lower) can be defined to differentiate between these QoS categories. This information is then used in the Dempster-Shafer framework to compute the probability that a source provides a report considered within a specific QoS category. An assessed source is considered trustworthy if the difference between the probabilities that its report belongs to the trustworthy category and it belonging to the untrustworthy category is greater than a threshold for trustworthiness. This approach also relies on reputation information in cases where direct experience with a source is insufficient.

To handle biases, Yu and Singh (2003) use a weighted majority algorithm. Weights are assigned to sources, which reflect their trustworthiness. Since their trust model is based on belief functions rather than scalars, the authors describe a method that maps belief functions to probabilities in order to compute the difference between a report and the observed trustworthiness, and accordingly updates the weights for each information source. Weights are initialised to 1.0 for every source, and tuned down after every interaction for which a source is regarded to provide inaccurate report. In this way, the effect of biased reports can be minimised over time in (fused) estimates. Since the ceiling is placed at 1.0, it would appear that weights of sources cannot increase with accurate information, but they decrease with unreliable reports.
Buchegger and Le Boudec (2004) propose an approach based on a Bayesian framework, with a prior distribution defined in the Beta family of distributions. A deviation test is used to classify information providers as trustworthy and untrustworthy, and reports that fail the deviation test are simply disregarded during fusion. This approach is oriented toward peer-to-peer and mobile ad hoc networks. A significant limitation of this work is the constant flooding of the network with reputational information even when this information is unsolicited or may not be useful to the recipient. Flooding the network with information may not be desirable in certain application contexts.

The majority of trust models adopt a subjectivist viewpoint (Goldstein, 2006). This is understandable, given that individuals have different satisfaction criteria or preferences. For instance, REGRET (Sabater and Sierra, 2001) allows agents to evaluate each other’s performance based on multiple domain dependent attributes (e.g., price or quality of service). Thus, while the criteria for good performance might be low price for one agent, timely delivery might be the defining factor for another. In terms of information, subjectivity may manifest in diverse perspectives, or different reporting patterns (Riloff et al., 2005). Problems arise when subjective views (of the evaluating agent) are used as grounds for assessing third-party opinions. This is especially the case in contexts where the motive is to determine objective facts about the state of the world. So far, none of the trust models discussed is equipped to deal with this problem: they disregard or tone-down opinions that do not align with the evaluator’s viewpoint. Different viewpoints may not be necessarily useless, especially when considering the benefits of complementarity of observations, and much sense may be gleaned from views that are seemingly adverse to personal opinion.

There do exist approaches better equipped to handle the problem of subjective opinions. For instance, POYRAZ (Şensoy et al., 2009) uses an ontology to describe each transaction. A source could exploit this mechanism to provide contextual information, for instance, on how an opinion was reached. A decision-maker can then decide for itself the relevance of the information to its decision-making. A possible drawback of this approach is the amount of meta-information that needs to be communicated along with the actual report. This may add an extra burden on communication media. Furthermore, information sources may not be willing to volunteer that much details, especially if doing so might reveal sensitive secrets (Dinur and Nissim, 2003). The other aspect to this is that there are limits to the extent to which an agent may locally aggregate and provide summary reports of multiple experiences. In cases where an information provider is known to consistently bias its reports (e.g., always exaggerating in a certain manner, or always reporting the opposite of the true situation (Gerardi et al., 2009), then the reports can be “aligned” (i.e., adjusted or re-interpreted based on observed pattern), thus reducing the need for discounting or discarding. This type of solution has been proposed by Regan et al. (2006) in their BLADE model, Teacy et al. (2012) in their HABIT model, and also by Koster (2012). More recently, Fang et al. (2014) use a two-layered clustering approach to categorise sources into different groups based on their subjectivity or reporting patterns. Reports from sources in different (subjectivity) groups are then aligned according to those of the evaluating agent.

In all the trust models discussed so far, a decision-maker must actively monitor the behaviour or the reliability of individual sources, and use this experience for the assessment of trust. It is also
likely that an agent may seek reputational information where the platform permits, and in situations where direct experience is lacking or insufficient. Situations exist, however, where direct and reputational information may be insufficient or unavailable (Burnett et al., 2010). For instance, the dynamic nature of a system may imply that a decision-maker does not gain sufficient experience to properly model the behaviour of sources. The same problem applies in large systems, where the frequency of interactions with any particular source may be very low. In addition, information about new-comers to the system may be relatively sparse, even in contexts of reputational opinion pooling. There are some interesting problems associated with this kind of setting, such as white-washing in open systems (Feldman et al., 2006), or assigning inappropriate weights to reports provided by unknown sources.

A few trust models have attempted to address this problem by exploiting correlations between behaviour and features of sources to form groups. The aim being that similar sources are grouped together. For instance, sources from the same organisation may be regarded as similar. Newcomers to the system may then be trusted based on the trust bestowed on other similar sources in the system. Examples of approaches in this category include Teacy et al. (2012), Burnett et al. (2010), and Liu et al. (2009).

We have adopted a similar idea of finding correlations between reports and source features to form groups, details of which is presented in Chapter 4. However, our motive and approach for group formation is quite distinct from those advanced in the above models. Among other things, source grouping allows us to manage the process of source selection.

2.4 Source Selection

We have discussed information fusion as a mechanism for combining observations (from different sources) to estimate the value of some environmental state. We have also discussed useful mechanisms, in trust and reputation, that serve the purposes of information transformation. In this section, we focus on a different aspect of multi-source integration: source selection. Source selection can be considered as a significant prerequisite to information fusion, where relevant information sources are integrated to meet a user’s information needs. This connection between source selection and fusion actually satisfies an important criterion of integration: that information provided by one module can serve another module, and ultimately enhance the operation of the overall system (Bozma and Duncan, 1994). The task of source selection involves incorporating the most appropriate source configuration or selection strategy from among the sources available to the system (Luo and Kay, 1989). A careful selection strategy not only has the potential to significantly improve fusion results, it can maximise the decision-maker’s confidence in the estimates (Becker and Corkill, 2007).

Deciding what sources to ask for opinions regarding the state of the world is a familiar problem. The underlying motives are also varied depending on the application domain and individual system requirements and/or constraints. For instance, Rowaihy et al. (2007) highlight the energy challenges of sensor networks and the need to prolong network lifetime as the reasons why the number of sources queried should be kept to a minimum. Zhang and Ji (2010) make a case for the selective integration of sources given the large number of available sources of information and the variability in their reliability. Doing so, the authors argue, will enable timely and efficient
decisions. The problem of information overload on the Web and social networks is examined by Heath (2008), with an emphasis on identifying trustworthy sources of relevant information to meet a user’s information needs. Information overload has also been identified as an issue in military contexts, and there are strong evidence that suggests that sensitive information do go unnoticed due to the large number of sources a decision-maker has to deal with. Šensoy et al. (2011) argue that only a subset of available sensing resources may be relevant for a specific task, and that the task constraints should guide decisions on what types of sources should and can be used. Tycho-giorgos and Bisdikian (2011) define information relevance as a measure of how “close” a piece of information is to what is desired by the user, and argue that the proliferation of potential information sources makes the process of selecting the most relevant sources challenging. Jøsang and Presti (2004) argue that the cost and risk analysis of a potential interaction may serve to inform whom to approach under what circumstances. Dong et al. (2012) argue that sometimes “less is more”, that the cost of integrating more sources may not be worthwhile if the benefit is limited. Lorenz et al. (2011) demonstrate how social influence can lead to correlated biases in information. In line with this, Berti-Equille et al. (2009) describe different kinds of correlated biases that may manifest in the source population, and argue that it is essential for a source selection mechanism to account for source dependency.

A large number of approaches exists for source selection, each serving a specific purpose. For instance, methods exist that concern with the maximisation of the quality of information obtained from the sources. Other approaches are focused on the minimisation of selection costs. Still, the focus of some is to balance the quality and cost trade-off. Different approaches affect truth discovery in different ways.

2.4.1 Crowd-Based Approaches

Crowd-based approaches are so called because they operate under the “wisdom of crowds effect”: the combination of as many sources as possible in order to minimise the influence of biased opinions (Surowiecki and Silverman, 2007). They do not explicitly reason about the number of sources to employ, or the cost of information acquisition. Most trust and reputation schemes fall under this category, their focus being solely to maximise the quality of information. Zhang et al. (2006) exploit the high density of nodes in sensor networks to estimate environmental states. Their model relies on the idea of the central limit theorem: that when multiple sources observe the same physical quantity independently, their reports will approximately follow a normal distribution. It is also often assumed, based on empirical rule, that about 68% of reports drawn from a normal distribution fall within one standard deviation of the true value. A decision-maker can filter out reports that fall outside a range, say, one standard deviation of the mean. The approach proposed by Whitby et al. (2004) is similar to that of Zhang et al. (2006). However, the model is based on a binomial distribution. Reports lying outside the $q$ or $1 - q$ quantile of majority reports are considered as outliers, and thus filtered out. Jianshu et al. (2006) also use filtering to deal with outlying reports, but, their approach is based on entropy. Entropy is a measure of uncertainty in information (Cover and Thomas, 2012). In Jianshu et al. (2006), if a new report leads to a significant change in the uncertainty in reports distribution, it is considered as biased and filtered out. Their specific methods for detecting outliers or biased reports may be different, but all the filtering approaches

---

2.4. Source Selection

have one thing in common: a reliance on large numbers of reports in order to obtain reasonable results in their statistical filtering.

Other approaches exist that do not necessarily filter out reports, but nevertheless rely on the crowd. Some examples are Wang et al. (2011), Teacy et al. (2006), and Aberer and Despotovic (2001). These approaches model the behaviours of individual sources, and discount their reports accordingly. Whereas the practice is often to delegate to the most reliable source in single-source settings (Ramchurn et al., 2004), similar to filtering approaches, these approaches do not employ an explicit mechanism for the selection of sources. One possible reason for this observation is perhaps as captured by Burnett (2011): “no agent rationally wishes to interact without gathering evidence about potential partners, but evidence can only be gathered through interactions”.

Of course the wisdom of crowd effect has stipulated conditions for the desired results to be obtained: the errors (or reports) of individual sources must be statistically independent (Vul and Pashler, 2008). As will be discussed later, this assumption often can’t be guaranteed. Also, it remains to be seen how crowd-based approaches adapt in environments with resource constraints; e.g., where resource use (i.e., querying a source) may affect the utility of a decision-maker.

2.4.2 Sampling-Based Approaches

In some domains, the selection of sources might be constrained in certain ways (e.g., cost, energy, timely decisions) (Wolfson et al., 2007). Approaches that operate under resource constraints often resort to the use of a subset of the source population. Even quality-driven approaches may, at some point, need to draw a line on the number of sources to integrate, even if they may not engage any active source selection strategy. This is especially the case in very large systems (Browne et al., 2007).

Generally speaking, source selection can be considered as a form of sampling: the process of selecting a subset of individuals from within a population to estimate some parameter of interest. Sampling is a popular statistical tool that has wide application in different areas including surveys. Advantages of sampling include reduced cost of information acquisition, greater speed of information gathering, greater scope, and greater accuracy (Cochran, 1977). Some of the traditional sampling techniques that form the basis for the development of more complex source selection strategies include (Lohr, 2009):

**Simple Random Sampling**

Simple random sampling is the simplest sampling technique. It involves selecting $n$ samples from the population $N$, such that every distinct sample has an equal probability of being selected. At any stage of the selection, each source not already selected is given an equal chance of being selected. This means that the $i$th sample had a marginally lower probability of being selected than the $(i+1)$th sample. Simple random sampling forms the basis for more complex sampling designs. For example, Przydatek et al. (2003) use this technique in the sensor domain to randomly select sensor nodes used to verify the correctness of fusion results.

**Stratified Random Sampling**

In stratified random sampling, the population of sources is divided into groups called strata. Random sampling is then applied independently to each of the strata in order to select the required candidates. The allocation strategy employed may be proportional, a case where a sampling fraction
in each of the stratum is proportional to that of the total population. An optimum allocation may also be employed, such that each stratum is proportionate to the variance of the distribution of the estimated variable. The subpopulations or subgroups in the different strata are non-overlapping, and together they comprise the whole of the population. The strata are often subgroups of interest to the decision-maker (e.g., sources owned by different organisations). Sources in the same stratum often tend to be more similar according to some measure of similarity. Therefore, an advantage of this technique is that it often increases precision of the estimate. This method has been applied to circuit-level power estimation (Ding et al., 1996).

**Cluster Sampling**

In cluster sampling, sources in the population are aggregated into larger groups or clusters. Each cluster should be as heterogenous as possible, but homogenous across clusters. Thus, each cluster is a scaled-down representation of the total population. Similar to stratified sampling, the subgroups should be both mutually exclusive and exhaustive. Subgroups are then selected at random. A variant of this technique is used in (Heinzelman et al., 2000) where sources are clustered into groups for data transmission purposes on a network.

A slightly different approach to employing a direct sampling technique is that proposed by Yu and Singh (2002). A decision-maker depends on a social network of acquaintances for referrals to “good” information sources. A referred source can further provide referrals, if it does not possess sufficient knowledge on the subject matter, thus creating a referral chain. Although this approach may end up selecting a desirable subset, the communication overheads, especially where there may be long referral chains, makes it very inefficient. To limit the resources expended in pursuing referrals, the authors define a depthLimit, a parameter that bounds the length of a referral chain. But the question still remains on what the appropriate depthLimit should be. As noted by the authors, shorter referral chains are likely to be more productive and accurate. This may be because longer referral chains introduce uncertainties in terms of the knowledge about the sources as one goes down the line. A similar source selection strategy is used in Yolum and Singh (2005).

One problem faced by sampling-based approaches is in determining appropriate numbers of sources to sample (Eckblad, 1991). Fortunately, various methods have been developed in statistics to determine an appropriate sample size given the desired accuracy and confidence (Cochran, 1977). For example, deviation bounds such as the Chernoff bound (Mitzenmacher and Upfal, 2005) have been applied in the context of trust to determine the minimum number of experiences necessary to achieve a desired level of confidence and error while estimating an agent’s reputation (Mui et al., 2002). While these bounds may be utilised by an agent to compute the potential number of sources to select, it is usually the case that this is an overestimate. This is because the bounds use parameters that are unknown in many situations, and, particularly, they are independent of the actual population size. This problem can be avoided if sampling is done in a more informed manner. For instance, the number of samples may be adjusted based on knowledge of their statistical properties (e.g., variance) (Watanabe, 2000). In line with this, Tran-Thanh et al. (2014) consider the problem of interdependent tasks allocation under budgetary constraints. In particular, given a budget, the aim of the approach is to determine the number of micro-tasks to be performed and the price to pay for each task. The authors use a quality control procedure known as AccurateAlloc, to efficiently allocate micro-tasks in order to minimise the probability of error.
2.4. Source Selection

The procedure employs a tuning or error bound parameter, such that only those candidate solutions that are within the error bound of the most popular candidate are progressed to a subsequent phase. While the proposed approach provides performance guarantees under fixed budget, its applicability might be limited in environments characterised by biased sources. This is because the authors assume that sources are not malicious, and therefore depend on basic consensus to predict correct answers. Thus, the number of sources required to attain a certain level of accuracy might increase dramatically depending on the number of ‘lazy’ workers in the system. Furthermore, the approach has a strong reliance on pre-assigned budgetary bounds for effective operation.

2.4.3 Decision-Theoretic Approaches

Decision-theoretic approaches regard source selection as an optimisation problem. By adopting a decision-theoretic view, a decision-maker can directly compute its expected utility, and use this to drive sampling decisions. The agent only pursues a course of action or chooses to integrate more sources if that is expected to lead to an increase in its utility.

Certain approaches under this category formulate the source selection problem as a Markov Decision Process (MDP). MDPs (Puterman, 2009) provide a sound mathematical framework for modelling problems of choice under uncertainty. The basic definition of an MDP is in terms of states, actions and rewards. The agent receives a reward by performing some action in a state. Using the reward model, the agent can make decisions on what actions to execute in order to maximise its long term reward or utility. Bernstein et al. (2003) use an MDP to perform adaptive peer selection on a peer-to-peer network, the goal being to select a viable peer from which a content may be downloaded. Their objective is to minimise the time it takes to obtain the content from a list of peers. Although the goal of this approach is to select a single peer for interaction, it provides a simple illustration of employing a decision-theoretic framework to guide the process of source selection. Castanon (1997) formulates the problem of dynamically scheduling a set of sensor resources for the classification of multiple targets as a Partially Observed Markov Decision Process (POMDP). Their approach, however, suffers from combinatorial explosion of the search space even in low-scale problems. The approach proposed by Kamar et al. (2012) uses an MDP to model a consensus task in a crowdsourcing application. At each point in time, the system must decide whether to incorporate an additional source or not, based on the likelihood of obtaining an accurate result. Hence, their action space is \{\text{hire, terminate}\}. If the suggested action is to hire, an additional source is selected from the pool of information sources available to the system, else the task terminates. The state space in their problem formulation consists of the complete history of sources’ reports that have been received up to the current time. Given that the state space grows exponentially, finding an optimal solution is infeasible even in small-horizon tasks. To deal with this complexity issue, the authors resort to sampling the solution space. In particular, they employ Monte-Carlo sampling to plan ahead and approximate the Value of Information (VoI). This is necessary in order to avoid sub-optimal source selection strategies, or, more specifically, to avoid wrong decisions of whether to continue or terminate incorporation of additional information source. However, long lookahead planning does not necessarily perform well in dynamic and uncertain environments (Kristensen, 1997). Another shortcoming of this approach is that the decision-theoretic planner does not reason about the reliability of potential sources. Thus, given the way the actions in the system are defined, both reliable and unreliable sources can be selected.
without much control of the learning process. The approach described by Teacy et al. (2008) operates in the context of a computational trust model, and therefore explicitly reasons about the reliability or trustworthiness of sources. An action space is defined in the model as all possible combinations of sources available in the system. At each point in time, a decision-maker maintains a belief state over the trustworthiness of the sources in each subset. Based on this belief state, the agent decides what action to perform (subset of sources to query). In particular, the authors use an approximation algorithm known as VPI (Value of Perfect Information) exploration method to estimate the value of selecting a particular subset of sources (or action). This approach, and a similar POMDP formulation in Irissappane et al. (2014) suffer from the same problems faced by the other MDP-based solutions already discussed. For instance, in a system having \( n = 100 \) sources, the learning algorithm in Teacy et al. (2008) would have \( 2^n \) different actions to select from in each belief state. Both the approach of Teacy et al. (2008) and the one described by Irissappane et al. (2014) have only been evaluated in relatively very small environments (with approximately \( \leq 40 \) sources).

Another view to the decision-making problem of source selection is to formulate it as a multi-arm bandit problem (Bergemann and Välimäki, 2006; Gittins, 1979). Multi-armed bandit (MAB) techniques have been extensively studied in statistics, with an equal interest in the areas of artificial intelligence, such as reinforcement learning (Sutton and Barto, 1998). A detailed treatment on this subject is provided in Chapter 3. However, the objective of the decision-maker is to allocate trials among \( K \) arms or alternatives in order to maximise its reward. An action is equivalent to a source selection strategy. The agent must balance the trade-off between exploitation of known actions or selection strategies and exploration of alternative action choices. Many algorithms have been proposed for this (Auer et al., 2002). A major difference between MAB and the full MDP-based problems (e.g., reinforcement learning) is that MAB is often assumed a single state process, whereas having multiple states is common in MDPs. The decision-maker doesn’t have to worry about state transitions in MAB, but need only concentrate on selecting actions that lead to better rewards. Pandey et al. (2007) exploit dependencies among arms to address the problem of high-dimensional MAB problems. Lu et al. (2010) describe an approach based on a contextual MAB, where side information or some domain knowledge is employed to aid the selection of an action. The reward in this case depends on both the action selected and the context. An example application where this may be used is in sponsored web search, for maximising the click-through rate (CTR) of an advert. Similarly, Li et al. (2010) use contextual MAB to recommend news sources to users based on contextual information about the users and news articles. Of course, by not explicitly modelling system states, MAB-based approaches lack the expressiveness of full MDPs. For instance, different actions might lead to different outcomes under different conditions or states. Such knowledge could aid the agent to explore in a more efficient manner. Traditional MAB problems deal with cases where only an arm may be played in each sampling round. However, many-real world applications has a combinatorial nature. Therefore, instead of selecting a single arm, a set of arms may be mapped to a single action. This variation to the classical multi-armed bandit formulation is referred to as combinatorial multi-armed bandit (Gai et al., 2010; Chen et al., 2013).

Active learning (Settles, 2009) is another related area under this source selection category.
2.4. Source Selection

The goal is typically to obtain more samples from a distribution of which a decision-maker is more uncertain. Thus, the agent’s decision to draw more samples is influenced by the variability in the distribution of the samples. Active learning has been applied within the contexts of stratified sampling for optimum allocation (Antos et al., 2010; Carpentier and Munos, 2011; Etoré and Jourdain, 2010) and multi-armed bandits (Antos et al., 2008). In the case of stratified sampling, the problem faced by the decision-maker is deciding the number of samples from each stratum. A typical approach is for the agent to sample each stratum and learn its distribution. The agent then uses this knowledge to allocate resources to the different strata. For example, consider a polling organisation that has to estimate as accurately as possible, within budgetary constraints, the number of people that are likely to vote “YES” in the Scottish Independence Referendum of 2014. In other words, the goal of the organisation is to improve the estimate with a limited set of samples. The organisation needs to concentrate on picking samples (individuals to ask) that represent the population well. If it has some side information or background knowledge about the different population groups (culture, ethnicity, location, religion, age, etc.), then this knowledge can be exploited to stratify the population. For instance, it might be known that people of a specific age group would mostly favour the “YES” vote. In addition, there might be evidence suggesting people resident in a certain region would be more inclined towards a “NO” vote. This information can provide more options for stratification. Budget can then be allocated to the different strata or subgroups identified. Whereas a naïve approach would be to allocate budget proportionally to the number of individuals in each group, if the survey organisation knew the level of variability within each group, it could make a more efficient allocation. Specifically, more individuals would be sampled from groups (e.g., geographical locations) with high variability in their voting preference.

Active learning attempts to estimate the variability, or variance, in a distribution, and use that as the basis for source selection. While this technique is quite efficient in terms of optimal allocation of sampling resources and precision in the estimates, prior knowledge of subgroups in the population, or of stratification variables is necessary (Podgurski et al., 1999). In addition, the active learner’s goal is that of variance reduction in the reports, and this is often used as a performance index for an algorithm. This can be problematic in situations where sources are biased. By not taking the reliability of groups into account, more sampling effort may be appropriated to unreliable groups, which may in turn impact truth discovery (Qi et al., 2013). Also, it is often the case in both MAB and active learning that the budget is set in advance. As argued in Reches et al. (2013), in certain contexts, an explicit budgetary bound may not be imposed, rather the onus is on the decision-maker to decide optimal allocations. This possibility is explored in this research, and is the focus of our sampling decision model presented in Chapter 5.

Other works worth mentioning under this section include Osborne et al. (2008). Their model for active data selection is based on a multi-output Gaussian process. The authors use Gaussian processes to represent correlations and delays between sensors, which then serve to inform when extra observations should be made, and from which sensors. While the approach they describe is grounded on sound mathematical principles such as Bayesian Monte Carlo method. Selection decisions are driven mainly by the uncertainty in data items rather than on knowledge of the sources. While there is nothing wrong with such a decision model, in some settings more flexibility might be required. For instance, their mechanism is unable to detect failed or unreliable sensors within
the population. This might lead to scarce resources (e.g., bandwidth) being used in querying such sensors, which in other circumstances would have been avoided. The approach described by Garnett et al. (2010), also uses Gaussian processes to perform inference on how to select subsets of sources for sensing tasks. Similar to Osborne et al. (2008), the approach described in this work does not explicitly account for the behaviour of sources, which in certain contexts, as earlier mentioned, might prove detrimental. This is especially the case where such background information about the behaviour of sources is available, or can be easily learned to enhance the process of source selection.

2.4.4 Dealing with Dependencies

The majority of source selection strategies we have discussed assume that information sources are independent. However, available evidence suggests otherwise (Lorenz et al., 2011; Tang et al., 2009; Anagnostopoulos et al., 2008; DeMarzo et al., 2003). For example, social influence can cause individuals to revise their personal observations (before sharing) (Lorenz et al., 2011). It is also possible that a (possibly misleading) report can be spread among large group of sources through copying (e.g., on the Web). An established statistical fact is that the integration of multiple sources or the wisdom of crowds effect works, if reports of individual sources are statistically independent (Lorenz et al., 2011; Vul and Pashler, 2008). Source dependence poses significant threats to the process of truth discovery, including the waste of resources (time, effort, money, etc.) in sampling redundant sources. Becker and Corkill (2007) argue that incorrectly assuming source independence might be tolerable when the sources are very reliable. However, incorrect independence assumptions can have huge (cost and quality) impact in systems with mediocre sources.

Several authors have recognised the significance of this problem, and have proposed different techniques to deal with dependence between information sources. For instance, Uddin et al. (2012) present an algorithm for diversifying the selection of information sources. The authors assume a static and prior knowledge of a stratification metric: they form a dependence graph on Twitter based on the assumption that individuals “following” others are likely to be dependent on those sources. While the proposed metric is quite relevant in the specific system, such knowledge may not always be available in other domains. Dong et al. (2009a) use an iterative method to estimate the probability of dependence between sources. Their approach relies on knowledge of ground truth, and works on the assumption that sources providing the same false reports are likely to be dependent. However, this sort of assumption can be problematic in environments where ground truth is unavailable. A Hidden Markov model is proposed by Dong et al. (2009b) for detecting copying relationships between pairs of sources. By not adopting a global view to the problem, this approach remains vulnerable to the effect of more complex source dependencies. However, the authors subsequently improve upon this work to include a global copying detection (Dong et al., 2010). Qi et al. (2013) use a probabilistic approach to determine the degree of dependency between sources and group them accordingly. Girvan and Newman (2002) investigate community structures in different networks (e.g., social and biological networks), which capture the relationships between entities. For instance, a community on a social network might represent people with similar interests or background. The authors argue that being able to identify these communities could help an agent understand and exploit the information platforms more effectively. We share a
similar view, and in Chapter 4 discuss an approach for developing a similar data structure, which allows us to support a more robust source selection mechanism.

### 2.4.5 Diversity

While source dependence is generally considered as a limiting factor in the context of collective intelligence, the diversity among sources strikes a positive note. In one sense, diversity can be employed as a tool to overcome some of the hindrances (e.g., correlated biases) to the effective integration of multiple sources of information. More importantly though, diversity adds perspectives to the truth discovery process that would otherwise be absent (Surowiecki, 2004; Wanous and Youtz, 1986). As Bonabeau (2009) argues, decision-making that exploits evidence from multiple sources must strike the right balance between diversity and expertise. Therefore, while traditional source integration problems have mainly focused on aspects such as accuracy or relevance of information, recent research shows that diversity is another highly desirable component (He et al., 2012; Ziegler et al., 2005).

One setback of information systems that focus only on accuracy, is that the opportunity to engage richer information contexts go unnoticed. Page (2010) shows that diversity makes a complex information system much more adaptable to changing information needs. This is particularly important in the sort of domain considered in this research, where interactions with sources might be quite restricted (e.g., due to resource constraints). Therefore, in terms of making source selection choices, diversity would ordinarily provide a decision-maker with a whole range of alternatives. It may be possible, for example, to approach one group of sources rather than another because their characteristics align with the user’s current information needs. This is the sort of idea behind Web searches and many information retrieval systems (Agrawal et al., 2009; Radlinski and Dumais, 2006).

In this research we build on the concept of source diversification. In that sense, a more detail treatment on the subject will be provided in Chapter 4.

### 2.5 Summary

In this chapter, we have provided a survey of related work in multi-source integration. In particular, we focused our discussion on approaches to information fusion. We discussed how existing approaches address the issues of uncertainty and bias in information. In the context of addressing the problem of uncertainty, we reviewed existing relevant work in trust and reputation, and reflected on their role in the truth discovery process. In addition, we reviewed different approaches for source selection.

The major barriers to the effective integration of multiple sources, which we aim to address in this research, have been highlighted as dependencies among information sources and constraints in resources. Although, we have reviewed different approaches for dealing with these challenges, most of these approaches deal with a single aspect of the problem, or are less relevant to our problem domain. For instance, the work of Tran-Thanh et al. (2014) while effective in handling budget allocation issues, do not consider the problem of biases, which may be introduced due to source collusion and misleading reports. Existing trust and reputation mechanisms, the focus of which is to address the problem source reliability (e.g., Teacy et al. (2006)), often do not take the cost of generating a good estimate into consideration, or are applicable to problems of very small
scale (e.g., Irissappane et al. (2014)).

As mentioned, the focus of this research is to optimally sample a population of sources to estimate ground truth. The model we present in Chapters 4 and 5 attempts to overcome some of the challenges faced by current approaches. In this way, we relax important limiting assumptions underlying existing approaches: (i) the greater the number of reports acquired, the better the estimate; (ii) reports are independent; and (iii) acquiring reports is cost-free.
Chapter 3

Background

In Chapter 1, we introduced a framework which underpins the contributions of this research. Within this framework, we cluster information sources into groups of similar sources. A strategy is then sought to efficiently sample from these groups in order to accurately estimate the value of some variable of interest.

In this chapter, we provide a précis of the techniques that we exploit in building a concrete realisation of the framework. These techniques underpin the model presented in Chapters 4 and 5. We begin in Section 3.1 with an overview of subjective logic, a belief calculus relevant for evidence combination and belief representation. In Sections 3.2 and 3.3, we examine two machine learning techniques: decision trees and clustering, useful for the formation of groups. Finally, in Section 3.4, we discuss reinforcement learning, a decision-theoretic framework used to guide the process of identifying effective sampling strategies.

3.1 Subjective Logic

As discussed in Chapter 2, there exist a number of approaches for evidence combination in multi-agent systems. The majority of these mechanisms are Bayesian. Bayesian systems use probabilities to represent degrees of belief about propositions of interest. An advantage of Bayesian systems is that they provide a theoretically sound basis for combining evidence, and are widely applied in both fusion and trust.

Subjective logic (SL) (Jøsang, 2013) is a form of probabilistic logic, which allows an agent to express opinions about the truth of propositions as degrees of belief, disbelief and uncertainty. In standard logic, propositions are considered to be either true or false. However, one cannot always determine with absolute certainty whether a proposition about the world is true or false. A simple example that illustrates this is the toss of a coin. If we are completely certain that the coin is fair, then we can confidently assign a probability of 0.5 to each outcome, “heads” or “tails”, meaning that both outcomes are equally likely. However, if we have some doubt over the fairness of the coin, then we are uncertain as to how to assign the probabilities. By modelling this notion of uncertainty explicitly, we can represent such problems in more detail, thus overcoming the constraints imposed by standard logic. Subjective logic extends classical probabilistic logic by also expressing uncertainty about the probability values themselves. This allows an agent to reason, or make assessments in the presence of uncertain or incomplete evidence.
3.1.1 Dempster-Shafer Theory

SL is based on Dempster-Shafer evidence theory (DST) (Shafer, 1976). DST builds on two ideas: obtaining degrees of belief for a proposition from subjective probabilities for a related proposition, and a rule for combining such degrees of belief based on evidence from multiple (presumably independent) sources. Thus, DST attempts to capture second-order uncertainty about opinions based on evidence from different sources. We present an example that illustrates these ideas, which provides the context for our discussion of SL.

Example 1

To illustrate how degrees of belief for one proposition (e.g., “it is snowing”) may be obtained from probabilities for another proposition (e.g., source y is reliable), suppose we (or an agent, A) have subjective probabilities for the reliability of y. Let the probability that source y (e.g., a weather sensor) is reliable be 0.8, and the probability that the source is unreliable is 0.2. Suppose y reports that it is snowing. This claim, which is considered to be true if y is reliable, is not necessarily false if the source is unreliable. The single piece of evidence obtained from the source justifies a 0.8 degree of belief that it is snowing, but only a zero (not a 0.2) degree of belief that it is not snowing. Interestingly also, the zero does not imply we can claim with certainty that it is not snowing, as would normally be the case in classical probabilistic logic. It merely represents the lack of evidence supporting a counter claim of no snow. The 0.8 and the zero together constitute a belief function. The 0.2 is our degree of ignorance, and can be interpreted as the lack of evidence to support either belief or disbelief in the proposition.

Suppose, we also have a 0.8 subjective probability for the reliability of another source, say, z, and a 0.2 probability of unreliability. Suppose this source also reports (independent from the first source) that it is snowing. As these two events (i.e., source y is reliable and source z is reliable) are considered independent, we may multiply the probabilities of these events. The probability that both sources are reliable is 0.8 \times 0.8 = 0.64, and the probability that neither is reliable is 0.2 \times 0.2 = 0.04. The probability that at least one is reliable is 1 – 0.04 = 0.96. Since they both reported that it is snowing, at least one of them being reliable implies that it is snowing. Thus, we may assign this proposition a degree of belief of 0.96. The consensus, therefore, leads to a stronger belief than any individual source’s opinion.

This example reveals two important elements of DST. First, there is no causal relationship between a proposition and its complement. Therefore, lack of belief does not necessarily imply disbelief. Rather, the lack of evidence supporting any proposition reflects a state of ignorance or uncertainty. Second, there is an intuitive process for narrowing a proposition, in which the state of uncertainty is initially given much weight, and replaced with belief or disbelief as more evidence is obtained.

We will return to this example whenever it becomes relevant to our discussion. In what follows, we introduce the key concepts of Dempster-Shafer theory relevant to SL.

**Definition 1 (Frame of Discernment)** A frame of discernment, \( \Theta \), is the set of all possible system states, exactly one of which is assumed to be true at any point in time.
For instance in Example 1, we can define the events \( X = \text{“it is snowing”} \), and \( Y = \text{“it is not snowing”} \). Our frame of discernment would then be:

\[
\Theta = \{X, Y\}
\] (3.1)

The elements of \( \Theta \) are often referred to as atomic states. However, given our imperfect knowledge of the world, we will not always be able to determine what state we are in. In that case, it makes sense to consider non-atomic states as well, consisting of the union of a number of atomic states. For instance, from our illustration in Equation 3.1, the power set, \( 2^\Theta = \{\{X\}, \{Y\}, \{X,Y\}, \emptyset\} \), consists of all possible unions of the atomic states in \( \Theta \). The elements in the set \( 2^\Theta \) represent propositions concerning the actual state of the system.

We can assign belief mass to various states (or sub-states of an atomic state), based on the strength of our belief (e.g., by using evidence from different sources) that the state is true.

**Definition 2 (Belief Mass Assignment)** Let \( \Theta \) be a frame of discernment. A function \( m : 2^\Theta \rightarrow [0,1] \) is defined as a belief mass assignment when it satisfies the following two properties:

1. \( m(\emptyset) = 0 \)
2. \( \sum_{x \in 2^\Theta} m(x) = 1 \)

For any element \( x \in 2^\Theta \), \( m(x) \) is known as its belief mass. The belief mass on an atomic state \( x \in \Theta \) is interpreted as the belief that the state in question is true. Belief mass on a non-atomic state \( x \in 2^\Theta \) is interpreted as the belief that one of the atomic states it contains is true, but there is uncertainty about which of them is true.

DST provides rules for combining evidence. By using these rules (e.g., consensus rule), it is possible to obtain a new belief mass that reflects the combined influence of two belief mass distributions. The key difference between DST and SL lies in these rules for combining evidence. Jøsang argues that the aggregation and consensus rules in DST are not consistent with Bayes’ theorem, and result in counter-intuitive results when faced with highly conflicting bodies of evidence (see counter-intuitive example under Section 2.2.1). SL attempts to address these issues (Jøsang, 2013).

### 3.1.2 Opinion Representation

Subjective logic operates on a 3-dimensional metric called opinion. The opinion model in SL can be considered as an instance of applying Dempster-Shafer theory of evidence.

**Definition 3 (Opinion)** Let \( \Theta = \{x, \neg x\} \) be a frame of discernment or state space containing \( x \) and its complement \( \neg x \). An opinion over \( x \) is defined as a tuple, \( \omega_x \).

\[
\omega_x \equiv (b_x, d_x, u_x, a_x)
\]

where \( b_x + d_x + u_x = 1 \), and \( (b_x, d_x, u_x, a_x) \in [0,1] \) (3.2)

Definition 3 corresponds to the sum of the belief masses in a belief mass assignment of DST, which according to Definition 2 sums up to 1. The notations \( b_x \), \( d_x \) and \( u_x \) represent belief, disbelief and uncertainty respectively, about the proposition \( x \). The variable \( a_x \) denotes the base rate, and is the a
priori probability about the validity of the proposition in the absence of any evidence. The default value of $a_x$ is typically taken to be 0.5 (Jøsang, 2013), which means that before any evidence is acquired, both outcomes ($x, \neg x$) are considered equally likely\(^1\). As we obtain evidence, the uncertainty, $u_x$, and the effect of $a_x$ decrease.

Opinions are often considered subjective, and will therefore be identified with an opinion owner whenever relevant. Suppose we are interested in modelling subjective opinions received from different sources in Example 1. We may consider a particular source $y$, and the proposition “it is snowing” or simply $x$. Then the notation $\omega_y^x$ represents an opinion held by $y$ about the truth of the proposition $x$. Further, suppose we wish to reason about opinions received from different sources with respect to the environmental state. For this reason, we can use $\omega_{yx}^A$ to represent an opinion held by an agent $A$ about source $y$ with respect to the proposition $x$. In the context of the example, $\omega_{yx}^A$ captures the trust that $A$ has in $y$ regarding what the source says about $x$.

Opinions can be projected onto a 1-dimensional probability space by computing the probability expectation value.

**Definition 4 (Probability Expectation)** Let $\omega_x = \{b_x, d_x, u_x, a_x\}$ be an opinion about the truth or validity of a proposition $x$, then the probability expectation of $\omega_x$ is:

$$E(\omega_x) = b_x + a_x u_x$$

(3.3)

The probability expectation value can be used as a single metric or a simpler mechanism to quantify an agent’s conviction about a proposition, the truth of which the agent is uncertain. For instance, this mechanism is used both in trust and fusion as a means for ordering entities (e.g., service providers), and for making decisions. Various approaches for ordering subjective opinions may be used, such as:

1. The opinion with the greatest probability expectation, $E(\omega_x)$ is the strongest opinion.
2. The opinion with the lowest uncertainty, $u_x$ is the strongest opinion.
3. The opinion with the lowest relative atomicity, $a_x$ is the strongest opinion.

Opinions can be represented graphically using a triangle as shown in Figure 3.1. The top vertex of the triangle represents maximum uncertainty, the bottom left represents maximum disbelief, and the bottom right represents maximum belief. An opinion can be uniquely identified as a point inside the triangle. The base rate $a_x$, and the probability expectation value $E(\omega_x)$ are also represented as points on the triangle. The base axis, between the disbelief and belief corners, is called the probability axis. The probability axis represents the state of zero uncertainty, and is equivalent to the traditional probability model. The base rate is represented as a point on the probability axis. The line joining the uncertainty corner and the base rate is known as the director. The probability expectation point can be geometrically determined by drawing a projection from the opinion point parallel to the director onto the probability axis.

\(^1\)This is regarded as the state of total uncertainty, represented as $\{0, 0, 1, 0.5\}$, and is the least informative state of an opinion.
3.1. Subjective Logic

The distance between an opinion point and the probability axis can be interpreted as the degree of uncertainty, and opinions where \( u \geq 0 \) are referred to as uncertain opinions. Opinions situated on the probability axis, that is, having \( u = 0 \), are referred to as dogmatic opinions. Opinions situated in the disbelief or belief vertex of the triangle, that is, having \( b_1 = 1 \) or \( d_1 = 1 \), are known as absolute opinions. They represent situations where it is absolutely certain that a state is either true or false. An example opinion \( w_1 = \{0.5, 0.1, 0.4, 0.6\} \), with expectation value \( E(\omega) = 0.74 \) is shown in Figure 3.1.

3.1.3 The Beta Distribution

A binomial opinion about a proposition can be modelled by a beta distribution, which represents the probability of the proposition being true based on evidence observed in support of or against it. The beta distribution is denoted as Beta\( (p \mid \alpha, \beta) \), where \( \alpha \) and \( \beta \) are its two evidence parameters. Equation 3.4 defines the beta probability density function (PDF), expressed using the gamma function \( \Gamma \), where \( 0 \leq p \leq 1 \), and \( \alpha, \beta > 0 \). The expected value of Beta\( (p \mid \alpha, \beta) \) is given in Equation 3.5.

\[
\text{Beta}(p \mid \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1 - p)^{\beta-1}
\]

\[
= \frac{(\alpha + \beta - 1)!}{(\alpha - 1)!(\beta - 1)!}
\]

\[
E(p) = \frac{\alpha}{(\alpha + \beta)}
\]
3.1. **Subjective Logic**

The Beta distribution provides a means to form opinions based on available evidence. Opinions formed can be updated in light of new evidence. For instance, let $r_x$ denote the number of positive experiences regarding $x$, and $s_x$ denote the number of negative experiences. The pair $(r_x, s_x)$ provides a source of $\alpha$ and $\beta$ parameters of the beta distribution such that:

$$
\alpha = r_x + 2a_x \\
\beta = s_x + 2(1 - a_x)
$$  \hspace{1cm} (3.6)

The parameter, $a_x$ in Equation 3.6 is the base rate (see explanation under Definition 3). Equation 3.5 can be re-written, thus:

$$
E(p) = \frac{r_x + 2a_x}{r_x + s_x + 2}
$$  \hspace{1cm} (3.7)

The $\alpha$ and $\beta$ parameters of the beta distribution can be directly mapped to the belief, disbelief and uncertainty components of an opinion, $b_x$, $d_x$ and $u_x$, using the following correspondence:

$$
\alpha = 2b_x/u_x + 2a_x \\
\beta = 2d_x/u + 2(1 - a_x)
$$  \hspace{1cm} (3.8)

The Beta distribution assumes different shapes with different amounts of evidence. For instance, when no evidence is present, that is $r_x = s_x = 0$, the a priori distribution, with default base rate $a_x = 0.5$ is the uniform beta PDF, with $\alpha = 1$ and $\beta = 1$ illustrated in Figure 3.2.

The case where $\alpha = \beta$ yields a PDF symmetric about 0.5 (the probability expectation value) as illustrated in Figure 3.3. Intuitively, this is the case when both outcomes are equally likely.

In Figure 3.4, the presence of more evidence leads to a more clustered distribution about the probability expectation value. Comparing Figures 3.3 and 3.4 gives an indication that the confidence in the expected value increases with more evidence. Therefore, the expected value is
3.1. Subjective Logic

more likely in the instance represented by Figure 3.3 than that of Figure 3.4.

![Figure 3.3: Symmetric PDF Beta(p | 5, 5)](image)

The illustration in Figure 3.5 shows how the distribution is skewed with more evidence in support of a proposition than against it (i.e., $\alpha > \beta$). The probability expectation value in the figure, according to Equation 3.5 is $E(p) = 0.8$. This value can be interpreted as saying that the relative frequency of a positive outcome in the future is somewhat uncertain, and that the most likely value is 0.8. This is the interpretation, and in fact the approach adopted by many trust models (e.g., Jøsang and Ismail, 2002; Teacy et al., 2006).

3.1.4 Evidence Aggregation

We have already seen how the parameters of the Beta distribution may be expressed in terms of the amount of positive and negative evidence regarding propositions. In a similar way, subjective logic provides tools for building opinions from evidence about propositions. As before, the amount of positive and negative evidence about a proposition $x$, can be represented as the pair $\langle r_x, s_x \rangle$. Evidence can be gathered in a number of different ways depending on the application.
context. For example, in online service selection, the outcome of a transaction (i.e., whether a transaction was successful or not) can be used as evidence about a proposition such as “seller X is trustworthy” (Jøsang et al., 2007). Also, in our running example, an agent may aggregate evidence from multiple sources regarding the proposition “it is snowing”. Reports from sources suggesting the presence of snow can be considered positive evidence about the proposition. Conversely, reports suggesting the absence of snow can be considered negative evidence in support of the proposition. The belief $b_x$, disbelief $d_x$, and uncertainty $u_x$ components of an opinion (see Definition 3) can then be expressed in terms of available evidence (i.e., positive $r_x$ and negative $s_x$ evidence):

$$b_x = \frac{r_x}{r_x + s_x + 2}$$
$$d_x = \frac{s_x}{r_x + s_x + 2}$$
$$u_x = \frac{2}{r_x + s_x + 2}$$  \hspace{1cm} (3.9)

From these definitions, we can re-define the probability expectation value in Equation 3.3 in terms of $\langle r_x, s_x \rangle$ in order to show the correspondence between binomial opinions and beta distributions (see Equations 3.5 and 3.6):

$$E(\omega_x) = \frac{r_x + 2a_x}{r_x + s_x + 2}$$  \hspace{1cm} (3.10)

Through the correspondence between binomial opinions and beta distributions, an algebra for subjective opinions has been defined. We briefly describe two of these operators, which are widely used in trust and fusion for combining and discounting opinions. We refer the reader to (Jøsang, 2013) for detailed description and proof of soundness of the algebra defined by SL.
3.1.5 SL Operators

Two (or more) opinions can be combined in SL to arrive at a new opinion. The consensus between two, possibly conflicting and uncertain, opinions is an opinion that reflects both opinions in a fair and equal manner. For instance, if two sources $y$ and $z$ have observed the outcomes of an event $x$ in two different conditions (e.g., spatial locations), and have formed two independent opinions about the likelihood of one outcome, then the consensus opinion is the belief about that outcome occurring, which a single agent would have after having observed $x$ in both conditions.

**Definition 5 (Consensus)** Let $\omega^y_x = (b^y_x, d^y_x, u^y_x, a^y_x)$ and $\omega^z_x = (b^z_x, d^z_x, u^z_x, a^z_x)$ be opinions respectively held by source $y$ and source $z$ about proposition $x$, and let $k = u^y_x + u^z_x - u^y_x u^z_x$. A consensus opinion, $\omega^{yz}_x$, between $\omega^y_x$ and $\omega^z_x$ is:

for $k \neq 0$

\[
\begin{align*}
    b^{yz}_x &= \frac{b^y_x u^z_x + b^z_x u^y_x}{k} \\
    d^{yz}_x &= \frac{d^y_x u^z_x + d^z_x u^y_x}{k} \\
    u^{yz}_x &= \frac{(u^y_x u^z_x)^2}{k} \\
    a^{yz}_x &= \frac{a^y_x u^z_x a^z_x u^y_x - (a^y_x + a^z_x) u^y_x u^z_x}{u^y_x u^z_x - 2u^y_x a^z_x}
\end{align*}
\]

for $k = 0$

\[
\begin{align*}
    b^{yz}_x &= \frac{b^y_x + b^z_x}{2} \\
    d^{yz}_x &= \frac{d^y_x + d^z_x}{2} \\
    u^{yz}_x &= 0 \\
    a^{yz}_x &= \frac{a^y_x + a^z_x}{2}
\end{align*}
\]

We use the symbol ‘$\oplus$’ to designate this operator, thus $\omega^{yz}_x \equiv \omega^y_x \oplus \omega^z_x$.

In discounting, it is assumed that an agent $A$ has an opinion about a source $y$. In addition, $y$ has an opinion about a proposition $x$ that it has shared with $A$. The agent can then form an opinion about $x$ by discounting $y$’s opinion about $x$ with its opinion about $y$.

**Definition 6 (Discounting)** Let $A$ and $y$ be two entities where $\omega^A_y = (b^A_y, d^A_y, u^A_y, a^A_y)$ is $A$’s opinion about $y$ as an information provider, and let $x$ be a proposition where $\omega^y_x = (b^y_x, d^y_x, u^y_x, a^y_x)$ is $y$’s opinion about $x$ expressed in a report to $A$. Let $\omega^{A,y}_x = (b^{A,y}_x, d^{A,y}_x, u^{A,y}_x, a^{A,y}_x)$ be the opinion such that:

1. $b^{A,y}_x = b^A_y b^y_x$
2. $d^{A,y}_x = d^A_y d^y_x$
3.2 Decision Tree Learning

A key challenge in machine learning classification, is the task of identifying which of a set of categories, classes or groups an object belongs. Classification is an instance of supervised learning: a form of learning that uses a set of trained or labelled examples to construct a model. In this section, we describe decision tree learning, a useful technique for discovering hidden patterns in data. Our main references on this subject are Breiman et al. (1984), Witten and Frank (2005), and Russell and Norvig (2010).

Decision tree learning (DTL) uses a tree structure as a predictive model for classification. The goal is to create a model (or learn a function) that predicts the value of a target (output) variable based on several input variables or features. The performance of the model is generally measured by the accuracy with which it classifies unseen cases. The input and output variables can be discrete or continuous. Trees having real-valued outputs are referred to as regression trees. Learned trees can also be represented as (if-then) rules to improve human readability. DTL has been successfully applied to a broad range of tasks including medical diagnosis and credit risk assessments.

![Decision Tree Diagram]

Figure 3.6: Example decision tree for credit rating

3. \( u_x^{A,y} = d_x^A + u_x^y + b_x^A u_x^y \)

4. \( a_x^{A,y} = a_x^y \)

Then \( \omega_x^{A,y} \) is called the discounting of \( \omega_x^y \) by \( \omega_x^A \) expressing A’s opinion about \( x \) as a result of \( y \)’s report to A. We use the symbol ‘\( \otimes \)’ to designate this operator, thus \( \omega_x^{A,y} = \omega_x^A \otimes \omega_x^y \).
### 3.2. Decision Tree Learning

#### 3.2.1 Decision Tree Representation

A decision tree is a tree structure, where nodes are either leaf nodes or decision nodes. A leaf node indicates the value of the target class of examples. A decision node specifies some test to be performed on some attribute of an instance. A decision tree classifies an instance by sorting it down the tree from the root to a leaf node. This process provides rules for classifying the instance. Figure 3.6 illustrates an example of a tree for credit rating. The tree classifies an individual as having a positive or negative credit rating, using a set of rules. A new case such as: \((\text{Income} = 45000 \text{ units}, \text{Credit history} = \text{Good})\), for example, would be sorted down the leftmost branch of the tree, and be classified as a positive credit rating. This decision tree may be expressed as a set of rules:

- If \(\text{Income} \leq 50000 \text{ units and Credit history} = \text{Good}\) then \(\text{Credit rating} = \text{Positive}\)
- If \(\text{Income} \leq 50000 \text{ units and Credit history} = \text{Bad}\) then \(\text{Credit rating} = \text{Negative}\)
- If \(\text{Income} > 50000\) then \(\text{Credit history} = \text{Positive}\)

#### 3.2.2 Constructing Decision Trees

The problem of constructing a decision tree can be expressed in a recursive manner. The basic idea involves selecting an attribute to place at the root node, and then making one branch for each possible value of the attribute. The process is repeated recursively for each branch, using only those instances that actually reach the branch. Development of a part of the tree is terminated if all instances at a node have the same classification.

Most algorithms for learning decision trees are based on a greedy divide-and-conquer strategy, where the most important attribute is tested first. The idea is to pick the attribute that goes as far as possible toward providing an exact classification of the examples. A perfect attribute divides the examples into sets, each containing only one class. To find a suitable candidate attribute, there must be a formal measure of the degree of “purity” at each node. The attribute that produces the purest daughter nodes is then chosen as the best attribute to split on. Before we discuss this procedure further, let’s consider the sample collection shown in Figure 3.1, which is used for the induction of the decision tree in Figure 3.6. There are 10 instances, and the class is labelled “Rating”. The purity at each node for both the Income and Credit history attributes can be
visualised in Figures 3.7 and 3.8, based on the training data in Figure 3.1. By inspecting the two tree stumps, we observe that *Income* is a better choice for splitting the tree, because it produces the purest daughter nodes. Next we define two measures used for this purpose. We then run through an example computation to illustrate the concept.

*Information gain* is normally used to decide which attribute to test at each node. This is calculated using a measure called *entropy*.

**Definition 7 (Entropy)** The entropy of a random variable $V$ with values $\text{Values}(V)$, such that $v \in \text{Values}(V)$, and each with probability $P(v)$, is:

$$
\text{Entropy}(V) = - \sum_{v \in \text{Values}(V)} P(v) \log_2 \frac{1}{P(v)} = - \sum_{v \in \text{Values}(V)} P(v) \log_2 P(v) \quad (3.11)
$$

Entropy is a measure of uncertainty of a random variable. For instance, a random variable with only one value (e.g., a biased coin that always turns up heads), has no uncertainty, and thus its entropy is zero. In other words, we gain no additional information by observing its value. A fair coin has an entropy of 1. In the context of decision tree learning, entropy characterises the degree of purity of an arbitrary collection of examples. Its measurement unit is *bits*.

**Definition 8 (Information gain)** Let $S$ denote a collection of examples, and let $A$ denote an arbitrary attribute. The information gain of attribute $A$ is the expected reduction in entropy caused by
3.2. Decision Tree Learning

Knowing the value of $A$.

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{S_v}{S} \cdot Entropy(S_v)$$  \hspace{1cm} (3.12)

In Equation 3.12, $Values(A)$ denotes the set of possibilities for values of attribute $A$. For a particular value, $v$ from the set $Values(A)$, $S_v$ denotes the set of examples that have value $v$ for attribute $A$.

We can apply this concept in order to evaluate the root of the $Income$ attribute in Figure 3.7. First, we calculate the entropy for the entire collection, $S$. The number of positive and negative classes are 6 and 4 respectively. $P(positive) = \frac{6}{10}$ and $P(negative) = \frac{4}{10}$. The entropy of $S$ can be calculated using Equation 3.11:

$$Entropy(S) = -(\frac{6}{10}) \log_2(\frac{6}{10}) - (\frac{4}{10}) \log_2(\frac{4}{10}) = 0.971 \text{ bits}$$  \hspace{1cm} (3.13)

Next, we compute the entropy for the various leaf nodes under $Income$. The number of positive and negative classes at the leaf nodes are [1, 4] and [5, 0] respectively.

$$Entropy(1, 4) = -(\frac{1}{5}) \log_2(\frac{1}{5}) - (\frac{4}{5}) \log_2(\frac{4}{5}) = 0.722 \text{ bits}$$

$$Entropy(5, 0) = -(\frac{5}{5}) \log_2(\frac{5}{5}) - (0) \log_2(0) = 0.0 \text{ bits}$$  \hspace{1cm} (3.14)

Finally, information gain according to Equation 3.12 can be derived from values obtained in Equations 3.13 and 3.14:

$$Gain(S, Income) = 0.971 - [(\frac{5}{10}) \times 0.722 + (\frac{5}{10}) \times 0] = 0.61 \text{ bits}$$  \hspace{1cm} (3.15)

In a similar manner, we can compute the information gain for $Credit$ history, resulting in $Gain(S, Credit history) = 0.124 \text{ bits}$. The result of the computation shows that $Income$ (with 0.61 bits) gains the most information, and therefore is a better fit for the root of the tree. In general, this process is conducted recursively to construct a decision tree. The process terminates when all leaf nodes are pure, or when the data cannot be split any further.

So far, we have described the decision tree algorithm at a generic level. However, there are many specific decision tree algorithms including ID3 (Quinlan, 1986), C4.5 (successor of ID3) (Quinlan, 1993), CART (Classification and Regression Trees) (Breiman et al., 1984), and M5 (Quinlan, 1992). Specific discussion of the variety of these algorithms is beyond the scope of this research. Rather, we focus on a class of algorithms concerned with numeric prediction, M5 being an example of this class.

3.2.3 Model Tree Learning

Classical tree induction methods are primarily focused on predicting the class to which a case belongs. There are, however, tasks that require the learned model to predict a numeric value associated with a case, rather than the class to which the case belongs. In what follows, we discuss model tree learning (Quinlan, 1992; Witten and Frank, 2005).
While leaves of classical decision trees are class labels, the leaves of a model tree are linear regression models. These models are used to estimate the target value. To induce a model tree, a splitting criterion is required. This criterion determines which attribute is the best to split the subset of the training data that reaches a particular node. Usually, the standard deviation, \( sd \) of the class values in the subset is used as a measure of the error at the node. The expected reduction in error is then calculated as a result of testing each attribute at that node. The attribute that maximises the expected error reduction is chosen for splitting at the node. The expected error reduction is:

\[
ER = sd(S) - \sum_i \frac{|S_i|}{|S|} \times sd(S_i),
\]

where \( S \) is the training set, and \( S_i \) is the set of instances that results from splitting the node according to the chosen attribute. The splitting process terminates when either of two conditions are met. First, when the class values of the instances that reach a node vary very slightly. Usually when their standard deviation is only, say, less than 5% of the standard deviation of the original instance set. Second, when just a few instances remain (e.g., four or fewer).

Once a tree is constructed, linear models are computed for each leaf node of the tree, using standard regression techniques. A smoothing process then attempts to compensate for any sharp discontinuities between the resulting linear models. Smoothing can be done by producing linear models for each internal node, as well as for the leaves at the time the tree is built. Typical smoothing calculation is:

\[
p' = \frac{mp + kq}{m + k},
\]

where \( p' \) is the prediction passed up to the next higher node, \( p \) is the prediction passed to the node from below, \( q \) is the value predicted by the model at the node, \( m \) is the number of training instances that reach the node below, and \( k \) is a smoothing constant.

Figure 3.9 shows an example model tree for CPU performance, with the leaf nodes as linear regression models. In the example model tree, computer configurations having value \( \text{MMAX} \leq 14000 \) are classified using linear model 1 (\( \text{LM1} \)), and those having \( \text{MMAX} > 14000 \) are classified using linear model 2 (\( \text{LM2} \)). In addition to the specific linear model for classifying
3.3 Cluster Analysis

Cluster analysis or clustering is the unsupervised classification of patterns (observations, data items, or feature vectors) into clusters or groups. In contrast with decision tree learning, clustering techniques are used when there is no class to be predicted, but rather when a set of objects are to be divided into natural groups. Unlike supervised learning, no pre-labelled instances are provided. The clustering problem can be formalised as follows: given a set of objects, characterised by some properties (e.g., features or attributes), group them in a meaningful way. An example of clustering is illustrated in Figure 3.11. The input space is shown in Figure 3.11(a), and the resulting clusters are shown in Figure 3.11(b). Points belonging to the same cluster are given the same label. In general, the goal of clustering is to place similar objects in the same group, and dissimilar objects in different groups. In this section, we discuss the general process for clustering. Our main references are (Jain et al., 1999) and (Everitt et al., 2011).

3.3.1 Components of Clustering

The task of clustering can be described in the following basic steps (Figure 3.12): feature selection and extraction; computing a similarity measure; and grouping. Feature selection involves identifying the most effective subset of the original features to use in clustering. Feature extraction is the use of one or more transformations of the input features to produce new salient features. Pattern representation refers to the number of classes, the number of available objects, and the number and type of the features available to the clustering algorithm. A similarity measure is some metric appropriate to the data domain that defines the proximity of pairs of objects. Grouping can be in terms of hard clustering, where each object belongs to a cluster or not, or soft clustering

<table>
<thead>
<tr>
<th>Linear Model 1 (LM 1)</th>
<th>Linear Model 2 (LM 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRP = 55.9</td>
<td>PRP = 10.2</td>
</tr>
<tr>
<td>+ 0.0489 x MYCT</td>
<td>+ 0.0041 x MYCT</td>
</tr>
<tr>
<td>+ 0.0153 x MMIN</td>
<td>+ 0.1351 x MMIN</td>
</tr>
<tr>
<td>+ 0.0056 x MMAX</td>
<td>+ 0.5022 x MMAX</td>
</tr>
<tr>
<td>+ 0.6410 x CACH</td>
<td>+ 0.1101 x CACH</td>
</tr>
<tr>
<td>- 0.2700 x CHMIN</td>
<td>- 0.890 x CHMIN</td>
</tr>
<tr>
<td>+ 1.480 x CHMAX</td>
<td>+ 2.965 x CHMAX</td>
</tr>
</tbody>
</table>

Figure 3.10: Linear models for CPU model tree
3.3. Clustering Analysis

Figure 3.11: Data clustering

![Data clustering diagram](image)

Figure 3.12: Stages in clustering

![Stages in clustering diagram](image)

(e.g., fuzzy clustering), where each object belongs to each cluster to a certain degree. The feedback loop involves cluster validity analysis, which is the assessment of the quality of the output (groups). This is the assessment of a clustering procedure’s output. The analysis uses a specific criterion of optimality, which is usually arrived at subjectively, depending on the objective of the application. In what follows, we discuss some of these stages in more detail.

3.3.2 Pattern Representation

Pattern representation involves the gathering of facts and conjectures about the data, and optionally performing feature selection and extraction.

**Definition 9 (Pattern)** A pattern (or object) $x$, is a single data item used by the clustering algorithm. It is usually represented by a vector of $d$ measurements:

$$x = \langle x_1, \ldots, x_d \rangle$$

**Definition 10 (Feature)** A feature (or attribute) $x_i$, is an individual scalar component of a pattern $x$. Each feature $x_i$, $i = 1, 2, \ldots, d$, can be considered a random variable. The variable, $d$, is the dimensionality of the pattern space.

Features can be either quantitative or qualitative. Quantitative features include:

(a) Continuous values (e.g., time)

(b) Discrete values (e.g., number of entities in a population)
3.3. CLUSTER ANALYSIS

(c) Interval values (e.g., duration of an event)

Qualitative features include:

(a) Nominal (e.g., gender)

(b) Ordinal (e.g., ratings such as “5-star”, “4-star”)

**Definition 11 (Pattern set)** A pattern set is denoted as $X = \{x_1, \ldots, x_n\}$. The $i$-th object in $X$ is:

$$x_i = (x_{i,1}, \ldots, x_{i,d})$$

The basic data for most applications of cluster analysis can be represented as $n \times d$ multivariate data matrix, $X$, containing the variable values describing each object to be clustered:

$$X = \begin{pmatrix}
    x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\
    x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\
    \vdots & \vdots & \ddots & \vdots \\
    x_{n,1} & x_{n,2} & \cdots & x_{n,d}
\end{pmatrix}$$

**Definition 12 (Class)** A class is a group of patterns having similar feature values according to a given metric. It is assumed that similar patterns are objects that most likely belong to the same distribution. Therefore, each class has its own model, which is typically mathematical in form; e.g., a Gaussian probability density function. A clustering algorithm attempts to group patterns so that the classes obtained reflect the true underlying distributions in the pattern set.

### 3.3.3 Similarity Measures

An important notion in the process of clustering is the similarity (or dissimilarity) between the objects to be arranged into groups. A measure of similarity is often carried out between two objects drawn from the same feature space.

**Definition 13 (Similarity)** Let $X$ be a set of objects, and $x_i, x_j \in X$. A similarity measure is defined as a function $\text{sim} : X \times X \rightarrow [0, 1]$ such that:

$$0 \leq \text{sim}(x_i, x_j) \leq 1$$

$$\text{sim}(x_i, x_i) = 1$$

$$\text{sim}(x_i, x_j) = \text{sim}(x_j, x_i)$$

$\text{sim}(x_i, x_j)$ defines the degree of similarity between the objects $x_i$ and $x_j$. The closer the objects in the feature space, the greater the similarity between them.

(Dis)Similarity between two patterns is usually computed using a *distance metric* defined on the feature space. The most commonly used measure for continuous features is the Minkowski metric:
When $p = 1$, $d_p$ specialises to the Manhattan distance, and when $p = 2$, it is known as the Euclidean distance.

One drawback of the Minkowski metric is the tendency of the largest-scaled feature (or features) to dominate. To resolve this issue, feature normalisation, or other weighting schemes can be performed as a pre-processing step.

While we have discussed only the most commonly used metrics here, other metrics are reported elsewhere. Ichino and Yaguchi (1994), for example, discuss metrics for computing similarity between objects with qualitative as well as quantitative features.

### 3.3.4 Clustering Algorithms

Clustering algorithms can be categorised into two main approaches: *partitioning* methods and *hierarchical* methods; see Figure 3.13.

In partitioning methods, the goal is to find $k$ clusters $C_1, C_2, \ldots, C_k$ of the input objects, that optimise a certain function. This function could be defined either locally (on a subset of the input objects), or globally (over all of the input objects). The most commonly used function is the squared error objective:

$$E = \sum_{i=1}^{k} \sum_{x \in C_i} d(x, m_i),$$

(3.19)

where $m_i$ is the centroid of the cluster $C_i$, and $d(x, m_i)$ is the distance between object $x$ and the centroid $m_i$. In general, the squared error clustering method attempts to minimise the distance
between each object and the centroid of the cluster to which the object is assigned. Thus, \( E \) in Equation 3.19 measures the total squared error incurred in representing the \( n \) objects, \( \mathbf{x}_1, \ldots, \mathbf{x}_n \), by the \( k \) cluster centers \( \mathbf{m}_1, \ldots, \mathbf{m}_k \). The optimal partitioning is one that minimises \( E \). The simplest and the most commonly used algorithm that uses a squared error objective is the \( k \)-means algorithm.

**\( k \)-means Clustering**

The \( k \)-means algorithm (MacQueen, 1967) starts with a random initial partition, and keeps reassigning objects to clusters based on the similarity between the objects and the cluster centers, until a convergence criterion is met. Usually, the process terminates when there is no reassignment from one cluster to another, or the squared error ceases to decrease significantly. The \( k \)-means algorithm can be accomplished in a series of steps as follows.

1. Choose \( k \) cluster centers
2. Assign each object to the closest cluster center
3. Recompute the cluster centers using the current assignment
4. If a convergence criteria is not met, go to step 2.

The \( k \)-means algorithm is easy to implement, with polynomial time complexity. A major drawback of this algorithm, is that it is sensitive to the selection of the initial partition. Thus the algorithm may converge to a local minimum of the criterion function, if the initial partition is poorly chosen. Also, while partitioning methods work well with large datasets, a general problem in using these methods is the choice of the number of clusters, which might not be known.

**Hierarchical Clustering**

In hierarchical clustering, the objects are divided into a series of partitions, which may vary from a single cluster containing all \( n \) objects, to \( n \) clusters each containing a single object. Hierarchical clustering may be subdivided into *agglomerative* methods and *divisive* methods. Agglomerative methods involve successively merging \( n \) individual objects into groups. Divisive methods on the other hand, involve separating a single group of \( n \) individual objects successively into finer groupings.

A hierarchical clustering algorithm yields a *dendrogram* representing the nested grouping of objects and similarity levels at which groupings change. An example dendrogram representing the partitioning of 7 objects into 4 groups is shown in Figure 3.14. The dotted line indicates the *cut-off point*, specifying the number of clusters required, or the desired dissimilarity level.

**Agglomerative Methods**

Agglomerative methods are perhaps the most widely used of the hierarchical methods. The general agglomerative algorithm involves the following steps:

1. Treat each object as a separate cluster, and compute the proximity matrix between each pair.
2. Merge the two most similar pair of clusters (given some proximity matrix) into a single cluster. Update the proximity matrix to reflect the merge operation.
3. If stopping criterion is satisfied, stop. Otherwise, go to step 2.
3.4 Reinforcement Learning

When faced with a problem of choice under uncertainty, an agent must learn to act in a *rational* way. In Section 3.1, we discussed how a single metric could be used to rank entities and for decision making. For example, when faced with the decision of buying from one of two service providers, an agent may decide to interact with the provider with the greatest expected probability of being reliable. Intuitively, interacting with a provider with the highest expectation value should provide a higher likelihood of a successful transaction outcome. Some decisions, however, involve complex choices, and employing a simple ranking metric, while useful in many settings, might not be sufficient. For example, what if each service commands a certain cost, controlled by (hidden) market forces? Or, what if there are other criteria or complex combinations of criteria other than, say, just reliability that is of concern to the agent? Ideally, an agent should be able to make

![An example dendrogram for a 7-object clustering](image)

**Figure 3.14:** An example dendrogram for a 7-object clustering

The proximity measure can be performed in a number of ways. The commonly used techniques are the *single-link* and *complete-link* algorithms.

In the single-link algorithm, the distance between two clusters is the minimum of the distances between all pairs of objects drawn from the two clusters. The linkage function is defined in Equation 3.20, where $C_1$ and $C_2$ are two clusters, and $d(x_1, x_2)$ denotes the distance (e.g., Euclidean distance) between the two objects, $x_1$ and $x_2$.

$$D(C_1, C_2) = \min_{x_1 \in C_1, x_2 \in C_2} d(x_1, x_2) \quad (3.20)$$

In the complete-link algorithm, the distance between two clusters is the maximum of all pairwise distances between objects in the two clusters. The linkage function is defined in Equation 3.21, where $C_1$ and $C_2$ are two clusters, and $d(x_1, x_2)$ denotes the distance between the two objects, $x_1$ and $x_2$.

$$D(C_1, C_2) = \max_{x_1 \in C_1, x_2 \in C_2} d(x_1, x_2) \quad (3.21)$$
decisions based on its current trade-offs as well as observations of its environment.

Decision theory focuses on the process of making rational decisions, and explicitly includes the pay-offs that may result. Central to decision theory is the notion of uncertainty about the domain or environment and utility to model the objective of the system. A decision-theoretic framework plays two roles. First, and most important, it provides a precise and concrete formulation of the problem to be solved. Second, it provides a guideline and method for designing a solution to the problem. In this section, we discuss reinforcement learning, a decision-theoretic framework that underpins the contributions of this research. While our discussion is kept at a general level, in Chapter 5 we will demonstrate how this technique is used to model sampling decisions. Our main references for reinforcement learning, and the fundamental solution methods are Sutton and Barto (1998), and Alpaydin (2004).

Reinforcement learning (RL) focuses on how an agent can achieve a goal by learning to behave in an approximately optimal way through trial and error interactions with its environment. The agent observes the state of the environment, selects an action to perform, and receives feedback or a reward from the environment. The goal of the agent is to maximise its cumulative reward in the long term.

Reinforcement learning is often referred to as “learning with a critic”, as opposed to “learning with a teacher”, which is the case with supervised learning. Rather than instructing a model by giving it correct actions, reinforcement learning uses evaluative feedback to guide the agent towards better action choices. The agent must be able to learn from its own experiences in the form of trial and error search to improve its performance over time. This kind of learning is best suited to interactive problems, as it is often impractical to obtain examples of desired behaviour that are both correct and representative of all the situations in which the agent has to act.

### 3.4.1 Bandit Problems

We begin by briefly discussing a simpler model in the context of the reinforcement learning problem, known as the $K$-armed bandit or multi-armed bandit (MAB). The $K$-armed bandit is a sequential decision problem modelled after a hypothetical slot machine with $K$ levers, or arms. Each arm when pulled provides a certain pay-off. The task is to decide which lever to pull to maximise the payoff. A gambler begins with no knowledge about the pay-off distributions of the arms, but through repeated play can make inferences about the true reward distributions of the arms. As the gambler is uncertain about the distributions, he must, at each iteration, decide between exploitation of the arm that has the highest expected pay-off according to its current knowledge, and exploration of alternatives. In exploring alternatives, the gambler can reduce the associated uncertainty and perhaps find an arm with a better pay-off.

Formally, the problem involves a slot machine with $K$ arms. At each time step $t = 1, 2, 3, \ldots$, one of the $K$ arms must be chosen to be played. Each arm $i$, when played, yields an immediate random real-valued pay-off or reward, according to some fixed (unknown) distribution. The random reward obtained from playing an arm repeatedly are independent and identically distributed, and independent of the plays of the other arms. An algorithm for the MAB problem must decide which arm to play at each time step $t$, based on the outcomes of the previous $t-1$ plays. Let $\mu_i$ denote the expected reward for arm $i$. Then the goal is to maximise the expected total reward in time $T$, i.e., $E[\sum_{t=1}^{T} \mu_i(t)]$, where $i(t)$ is the arm played in time step $t$, and the expectation is over
3.4. Reinforcement Learning

The classical bandit problem is equivalent to a one-state reinforcement learning problem. The full reinforcement learning problem generalises this one-state case in a number of ways. First, the system is associated with several states. The states could be likened to having several slot machines, each having its own set of arms with their respective pay-off distributions. Second, when an action is performed (an arm is pulled), it affects not only the pay-off (reward), but also the next state. Third, the reward could be delayed in some cases, and we should be able to estimate immediate values from delayed rewards.

3.4.2 Elements of Reinforcement Learning

In reinforcement learning, the agent interacts with its environment at each of a sequence of discrete time step, \( t = 1, 2, 3, \ldots \). At each time step \( t \), the agent observes the environmental state, \( s_t \in \mathcal{S} \), where \( \mathcal{S} \) is the set of all possible states. On that basis, the agent selects and performs an action \( a_t \in \mathcal{A} \), where \( \mathcal{A} \) is the set of all possible actions. As a consequence of its action \( a_t \) in state \( s_t \), the agent receives a numerical reward \( r_{t+1} \in \mathbb{R} \) one time step later, and the system moves to the next state \( s_{t+1} \in \mathcal{S} \).

The reward and next state are sampled from their respective probability distributions, \( \mathbb{P}(s_{t+1}|s_t, a_t) \) and \( \mathbb{P}(r_{t+1}|s_t, a_t) \). The expression \( \mathbb{P}(s_{t+1}|s_t, a_t) \) is the probability of reaching state \( s_{t+1} \) when action \( a_t \) is taken in state \( s_t \). Similarly, \( \mathbb{P}(r_{t+1}|s_t, a_t) \) is the probability of getting reward \( r_{t+1} \) when action \( a_t \) is taken in state \( s_t \).

The domain of the reinforcement learning problem is modelled as a Markov Decision Process (MDP) with parameters \( \langle \mathcal{S}, \mathcal{A}, \mathbb{P}_s, \mathbb{P}_r \rangle \). The Markov property entails that the state and the reward in the next time step, \( s_{t+1} \) and \( r_{t+1} \), depend only on the current state, \( s_t \) and action, \( a_t \).

Policies

A policy, \( \pi \), is a mapping from perceived states to actions: \( \pi : \mathcal{S} \rightarrow \mathcal{A} \). The policy defines the agent’s behaviour or the action to be taken in any state \( s_t : a_t = \pi(s_t) \). A stationary policy specifies the same action each time a state is visited. A stochastic policy specifies an action independently from the same probability distribution over the possible actions each time a state is visited. The agent’s behaviour often changes during learning, such that it neither leans towards a stationary or stochastic policy. However, once an optimal policy is found, the agent’s policy will likely become...
stationary. The process of solving a reinforcement learning problem involves finding a policy that maximises the long-term reward of the agent.

**Task Model and Reward Function**

The aim of the agent in a reinforcement learning problem is to maximise the reward it receives. The agent is not merely interested in maximising the immediate reward in the current state, but is interested in maximising its reward in the longer term. In this regard, we differentiate between two kinds of task models: episodic and continuous tasks.

In an episodic or finite-horizon task, the agent’s interaction with the environment naturally breaks down into a sequence of separate episodes. An episode, or trial, is the sequence of actions from the start to the terminal state. An episode ends when state $s_{t+1}$ is a final state or “absorbing state”. The agent’s goal in this context is to maximise the expected reward for the next bounded time steps $T$, where $T$ is the final time step. The total reward or return is given as:

$$R_t = r_{t+1} + r_{t+2} + \ldots + r_{t+T} = \sum_{k=0}^{T} r_{t+k+1}$$ (3.22)

This kind of model is useful in applications where there is a clearly defined termination point or final time step (e.g., game of chess). However, certain tasks are on-going, and there are no pre-defined termination points (at least as perceived by the agent).

In the continuous or infinite-horizon task, future rewards are discounted by a constant discount factor $\gamma$, $0 \leq \gamma < 1$:

$$R_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \ldots = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}$$ (3.23)

The discount factor $\gamma$, determines the value of future rewards, such that a reward received $k$ time steps in the future is worth only $\gamma^{k-1}$ times what it would be worth if it were received immediately. If $\gamma = 0$, then only the immediate reward counts. As $\gamma$ approaches 1, rewards are biased towards the future. For example, in a sensor network where the nodes are often constrained by battery life, one might prefer rewards sooner rather than later because of the uncertainty in the lifetime of the sensor nodes.

The reward function or objective function defines the goal of a reinforcement learning problem. The function maps each perceived state-action pair to a single numeric reward. The reward function specifies what the good and bad events are for the agent. This, as a matter of importance, may serve as a basis for changing the current policy. For example, if a current course of action leads to a lower reward, an agent may change the strategy in favour of one that guarantees a better reward.

Many real-life problems involve dealing with multiple objectives. For example, an information consumer may want to optimise both the cost of obtaining information (by sampling a group of sensors) and the quality of information (from the sensors). In such contexts, instead of getting a single reward signal, the agent receives a reward divided up into several components in the form of a reward vector. The reward signal $r(s,a)$ is decomposed into a vector: $\bar{r}(s,a) = [r_1(s,a), r_2(s,a), \ldots, r_z(s,a)]$, and an agent could potentially optimise many different functions of this reward (Coello et al., 2002; Gábor et al., 1998; Barrett and Narayanan, 2008).
Value Functions

The value function is closely related to the reward function. Unlike the reward function, which specifies the desirability of an action in the immediate sense, a value function specifies what is good in the long run. The value of a state $s_t$ under a policy $\pi$, is the expected cumulative reward that an agent would receive if the agent follows policy $\pi$, starting from state $s_t$:

$$V^\pi(s) = E_\pi \{ R_t | s_t = s \} \quad (3.24)$$

Similarly, the value of taking action $a_t$ when in state $s_t$ under policy $\pi$, is the expected cumulative reward that the agent would receive starting from state $s_t$, taking action $a_t$, and thereafter following the policy $\pi$:

$$Q^\pi(s, a) = E_\pi \{ R_t | s_t = s, a_t = a \} \quad (3.25)$$

The state value function $V^\pi(s_t)$ denotes how good it is for the agent to be in state $s_t$, whereas, the state action value $Q^\pi(s_t, a_t)$, denotes how good it is to perform action $a_t$ in state $s_t$.

Optimal Policies

An agent’s aim is to find a policy that is optimal, such that, starting from any state, following the policy yields the maximum possible expected reward that can be achieved. For each policy, $\pi$, there is a state value function $V^\pi(s_t)$, and the agent’s aim is to find the optimal policy $\pi^*$, such that:

$$\forall s_t, V^*(s_t) = \max_{\pi} V^\pi(s_t) \quad (3.26)$$

Similarly, for the state action pair $Q(s_t, a_t)$, we have:

$$\forall s_t, a_t, Q^*(s_t, a_t) = \max_{a_t} Q^\pi(s_t, a_t) \quad (3.27)$$

Once the $Q^*(s_t, a_t)$ values have been obtained, the agent can simply choose the action $a_t^*$, which has the highest value among all $Q^*(s_t, a_t)$, such that $Q^*(s_t, a_t^*) = \max_{a_t} Q^\pi(s_t, a_t)$.

3.4.3 Reinforcement Learning Methods

In reinforcement learning the learner is the decision-maker that takes actions in an environment, and as a consequence receives a reward (or punishment in some cases) in trying to solve a problem. Here we describe three fundamental and widely used methods for solving reinforcement learning problems. These learning methods focus on estimating value functions and finding optimal policies.

Model-Based Learning

In model-based learning, the learner has complete knowledge of the environment model parameters, $P_s$ (probability of reaching some state) and $P_r$ (probability of acquiring some reward). The optimal value function and policy can be directly solved using dynamic programming, without the need for any search or exploration. Dynamic programming (DP) (Bertsekas, 1995) refers to a
collection of techniques for efficiently solving a broad range of optimisation problems, and can be used to solve for optimal policies.

To find the optimal policy, the optimal value function can be used. There is an iterative algorithm provided in DP called *value iteration* (Algorithm 1), which has been shown to converge to the correct $V^\pi$ values. This method iteratively computes the state value function $V(s)$ with the maximum value taken over all actions. It then derives the required policy based on this optimal value function. The values are said to converge if the maximum value difference between two iterations is less than a certain threshold $\zeta$: $\max_{s \in S} |V^{(l+1)}(s) - V^{(l)}(s)| < \zeta$, where $l$ is the iteration counter.

\section*{Algorithm 1 \ Value iteration algorithm for model-based learning.}
\begin{algorithmic}
   \State Initialise $V(s)$ arbitrarily for all $s \in S$
   \Repeat
   \For all $s \in S$
   \For all $a \in A$
   \Statex $Q(s, a) \leftarrow E[r|s, a] + \gamma \sum_{s' \in S} P(s'|s, a)V(s')$
   \Statex $V(s) \leftarrow \max_a Q(s, a)$
   \EndFor
   \EndFor
   \Until $V(s)$ converge
\end{algorithmic}

Another widely used algorithm in model-based learning for finding the optimal policy is *policy iteration* (Algorithm 2). In policy iteration, the computation is carried out on the policy itself, rather than doing it indirectly over the values. That means, the policy rather than the values is stored and updated. The process involves a policy *evaluation* and *improvement* sequence. The idea is to start with a policy and improve it repeatedly until no improvement is guaranteed.

\section*{Algorithm 2 \ Policy iteration algorithm for model-based learning.}
\begin{algorithmic}
   \State Initialise $\pi'$ arbitrarily
   \Repeat
   \State $\pi \leftarrow \pi'$
   \State Compute the value using $\pi$ by solving the linear equation
   \Statex $V^\pi(s) = E[r|s, \pi(s)] + \gamma \sum_{s' \in S} P(s'|s, \pi(s))V^\pi(s')$
   \State Improve the policy at each state
   \Statex $\pi'(s) \leftarrow \arg\max_a (E[r|s, a] + \gamma \sum_{s' \in S} P(s'|s, a)V^\pi(s'))$
   \Until $\pi = \pi'$
\end{algorithmic}

Once a policy, $\pi$, has been improved using $V^\pi$ to obtain a better policy, $\pi'$, $V^{\pi'}$ can then be computed, and further improved to yield an even better policy, $\pi''$. We can thus obtain a sequence of monotonically improving policies and value functions as illustrated in Figure 3.16, where $\xrightarrow{E}$ denotes a policy evaluation stage, and $\xrightarrow{1}$ denotes a policy improvement stage. Each policy is guaranteed to be a strict improvement over the previous one (unless it is already optimal). As a
finite MDP has only a finite number of policies, this process must converge to an optimal policy and optimal value function in a finite number of iterations.

One drawback of policy iteration, is that it is known to be computationally more expensive than value iteration. Each iteration involves policy evaluation, which may itself be a protracted iterative computation requiring multiple iterations through the state set.

**Monte Carlo Methods**

Monte Carlo (MC) methods (Hammersley and Handscomb, 1964) are stochastic techniques, which rely on random components to estimate solutions to problems. In RL, MC methods are ways of solving the reinforcement learning problem based on averaging sample returns. MC methods, unlike DP, do not require complete knowledge of the environment. They require only experience, which in this context refers to sample sequences of states, actions, and rewards from an online or simulated interaction with the environment. It is assumed that experience is divided into episodes, and so, MC methods are defined only for episodic tasks. The value functions, \( V(s) \) or \( Q(s, a) \), are only updated upon the completion of an episode. MC methods are thus incremental in an episode-by-episode sense, not in a step-by-step sense.

A representative MC algorithm, based on the idea of policy iteration, is provided in Algorithm 3. Similar to policy iteration in model-based learning, the algorithm first evaluates the value function \( Q(s, a) \) under an arbitrary policy \( \pi_0 \). It then improves \( \pi_0 \) using \( Q(s, a) \) to yield a better policy \( \pi_1 \). This sequence of **evaluate-improve** then continues on an episode-by-episode basis, and the approximate policy and the approximate value function asymptotically approach their optima.

**Temporal-Difference Learning**

Temporal-difference learning (TD) (Tesauro, 1995) combines the ideas of DP and MC. Like Monte Carlo methods, TD can learn directly from experience without a model of the environment (agents seldom have perfect knowledge of the environment). Also, similar to DP, TD updates estimates based on other learned estimates, without the need to wait for a final outcome. TD thus provides an answer to the more interesting and realistic application of reinforcement learning: the situations where no model is assumed and the learning task is not constrained to episodes.

One of the most widely used TD or *model-free* algorithms is known as **Q-learning** (Watkins, 1989; Watkins and Dayan, 1992). For any state \( s \), Q-learning chooses an action \( a \) to perform such that the state-action value \( Q(s, a) \) is maximised. The algorithm uses a state-action value updating or running average strategy in order to estimate the optimal policy (Equation 3.28). This is particularly important since for the same state \( s \) and action \( a \), an agent may receive different rewards or move to different next states.

\[
\hat{Q}(s_t, a_t) \leftarrow \hat{Q}(s_t, a_t) + \eta(r_{t+1} + \gamma \max_{a_{t+1}} \hat{Q}(s_{t+1}, a_{t+1}) - \hat{Q}(s_t, a_t))
\]  

(3.28)
We can consider Equation 3.28 to be reducing the difference between the current \( Q \) value, and the backed-up estimate from one time step later. The \( r_{t+1} + \gamma \max_a Q(s_{t+1}, a) \) values can also be thought of as samples of instances for each \((s_t, a_t)\) pair, and the aim is for the estimate \( \hat{Q}(s, a) \), to converge to its mean. The learning rate, \( \eta \), is gradually decreased over time for convergence. This algorithm has been shown to converge to optimal \( Q^* \) values (Watkins and Dayan, 1992). The pseudocode of the Q-learning algorithm is given in Algorithm 4.

**Algorithm 4 Q-learning Temporal Difference algorithm.**

<table>
<thead>
<tr>
<th>Line</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Initialise ( Q(s, a) ) arbitrarily for all ( s \in S ) and all ( a \in A )</td>
</tr>
<tr>
<td>2</td>
<td>For all episodes</td>
</tr>
<tr>
<td>3</td>
<td>Initialise ( s )</td>
</tr>
<tr>
<td>4</td>
<td>Repeat</td>
</tr>
<tr>
<td>5</td>
<td>Choose an action ( a ) using policy derived from ( Q )</td>
</tr>
<tr>
<td>6</td>
<td>Take action ( a ), observe ( r ) and ( s' )</td>
</tr>
<tr>
<td>7</td>
<td>Update ( Q(s, a) ) :</td>
</tr>
<tr>
<td>8</td>
<td>( Q(s, a) \leftarrow Q(s, a) + \eta(r + \gamma \max_a Q(s', a') - Q(s, a)) )</td>
</tr>
<tr>
<td>9</td>
<td>( s \leftarrow s' )</td>
</tr>
<tr>
<td>10</td>
<td>Until ( s ) is terminal</td>
</tr>
</tbody>
</table>

Q-learning is regarded as an off-policy method. An off-policy learner learns the value of the optimal policy independently of the agent’s actions. This approach may prove risky, possibly leading to large negative rewards, as a result of not taking into account the costs associated with exploration.
An alternative is offered by way of *on-policy* learning. The on-policy method involves learning the policy ‘on the fly’ along with exploration steps, and with the current policy being used to determine the next action. The on-policy counterpart of Q-learning is the SARSA algorithm (Algorithm 5). The major difference between SARSA and Q-learning is that the maximum reward for the next state is not necessarily used for updating the Q-values. Instead, a new action (and therefore reward) is selected using the same policy that determined the original action. An experience in SARSA is encoded in the form $(s, a, r, s', a')$. This means that when the agent was in state $s$, it performed action $a$, received reward $r$, and ended up in state $s'$, from which it decided to perform action $a'$. This provides a new experience to update the state-action value $Q(s, a)$. The new value that this experience provides is obtained using $r + Q(s', a')$ (see Algorithm 5).

SARSA has better convergence guarantees in comparison to Q-learning. The algorithm learns much more rapidly, and its average policy is shown to be better than that of other RL approaches (Sutton and Barto, 1998).

**Algorithm 5** SARSA Temporal Difference algorithm.

<table>
<thead>
<tr>
<th>Initialise $Q(s, a)$ arbitrarily for all $s \in S$ and all $a \in A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>For all episodes</td>
</tr>
<tr>
<td>Initialise $s$</td>
</tr>
<tr>
<td>Choose an action $a$ using policy derived from $Q$</td>
</tr>
<tr>
<td>Repeat</td>
</tr>
<tr>
<td>Take action $a$, observe $r$ and $s'$</td>
</tr>
<tr>
<td>Choose an action $a'$ using policy derived from $Q$</td>
</tr>
<tr>
<td>Update $Q(s, a)$ :</td>
</tr>
<tr>
<td>$Q(s, a) \leftarrow Q(s, a) + \eta (r + \gamma Q(s', a') - Q(s, a))$</td>
</tr>
<tr>
<td>$s \leftarrow s'$, $a \leftarrow a'$</td>
</tr>
<tr>
<td>Until $s$ is terminal</td>
</tr>
</tbody>
</table>

### 3.4.4 Action Selection Strategy

In learning a policy, the agent must find a balance between exploiting current knowledge and exploring alternative (and possibly unknown) action choices. For example, to obtain a high reward, the agent must prefer actions tried in the past and known to produce good outcomes. To discover such actions, however, it has to try actions, of which it has little or no knowledge. This process is particularly significant in tasks with stochastic rewards. Each action must be tried numerous times to gain a reliable estimate of its expected reward. This is known as the exploration-exploitation dilemma, whereby the agent can neither explore nor exploit exclusively without failing at a task, the problem of which is well-studied (Gittins, 1979; Lai and Robbins, 1985; Vermorel and Mohri, 2005; Langford and Zhang, 2007).

To excel in its task, the agent needs sophisticated strategies for action selection. In other words, the agent needs a strategy that specifies what action to take in each situation. The aim of the strategy is to balance the trade-off between exploitation and exploration, by not always exploiting what has been learnt so far. We now describe three widely used strategies.
3.5 Summary

Greedy Strategy
The simplest action selection strategy is to select the action with the highest estimated action (or state-action) value. That is, when in state $s$, an agent selects an action $a^*$, such that:

$$Q(s,a^*) = \max_a Q(s,a).$$  \hspace{1cm} (3.29)

$Q(s,a)$ is the current estimate of the true value $Q^*(s,a)$ of action $a$ in state $s$. This strategy always exploits current knowledge to maximise the agent’s immediate reward. Seemingly inferior actions are never explored to see if they might actually be better.

$\epsilon$-greedy Strategy
A slightly different strategy is to behave greedily most of the time, and randomly at other times. The $\epsilon$-greedy strategy normally selects the action with the highest estimated reward to perform, but with a small probability $\epsilon$, an action is selected at random, uniformly, and independent of the state-action value estimates.

Softmax Strategy
One disadvantage of the $\epsilon$-greedy strategy is that (alternative) random actions are selected uniformly. The worst action is just as likely to be selected as the second best action. The softmax (e.g., Boltzmann exploration) strategy deals with this problem in a more principled manner. Boltzmann exploration selects an action $a$ to execute in state $s$ with probability proportional to its average reward tempered by $T_{mp}$, which decreases over time to favour exploration earlier in the learning process:

$$P(a|s) = \frac{e^{Q(s,a)/T_{mp}}}{\sum_{a \in A} e^{Q(s,a)/T_{mp}}}$$  \hspace{1cm} (3.30)

High $T_{mp}$ values cause all the actions to have almost the same probability of being selected. Low temperature values, on the other hand, differentiate more between action selection probabilities, given their value estimates. Softmax is a good strategy to employ in situations where the worst actions are very unfavourable. As noted by Sutton and Barto (1998), it is not clear which of these strategies provides the best results overall. In line with the “no free lunch theorem” (Wolpert and Macready, 1997), the cost of finding a solution averaged over all problems in a class may well be the same for any strategy adopted. The nature of the task will have some bearing on how well each strategy influences learning.

3.5 Summary
In this chapter, we have presented details of a number of techniques that we exploit within our model. We have introduced subjective logic and reinforcement learning, which we will use in subsequent chapters for evidence combination, and in guiding sampling decisions. In addition, we discussed two machine learning techniques: decision trees and clustering, that are suitable for the process of forming groups.
Chapter 4

Source Diversification

In Chapter 2, we outlined existing approaches to source selection and fusion, and highlighted some of the problems posed by resource-constrained environments. In this chapter, we argue that agents operating in complex, dynamic and constrained environments can adopt a model of diversity in order to minimise redundant sampling and mitigate the effect of certain biases. Exploiting source diversity may, for example, provide evidence from groups of sources with different perspectives on a problem. This has the potential to mitigate the risk of double-counting evidence due to correlated biases among group members. We adopt techniques from machine learning in order to identify complex behaviour patterns, and to disambiguate what metrics lead to a good stratification of the source population. Where relationships exist between sources’ features and their reports, a model of diversity can help a decision-maker to avoid redundant sampling, and to make better assessments. To support our argument, we present and evaluate our model of diversity in the context of a simulated multi-agent system involving a population of information sources and a decision-maker. Using a simulated environment allows us to explore the performance of our model more broadly than a specific dataset would allow us to. However, we acknowledge that our approach could be evaluated within the contexts of Weather Underground \(^1\) and other sources for data. In what follows, we adapt our example from Chapter 3 to illustrate the significance of diversity in the context of our problem. This will set the basis for our discussion in the rest of the chapter.

Example 2

Suppose agent A has access to four sources, the reports from whom can be used to make weather predictions. Each source has a 0.8 probability of being accurate, this probability being estimated from observations of past performance. Suppose that the current task faced by the agent is to predict whether or not it is going to snow. How can the reports from these sources be integrated efficiently to obtain the most accurate possible result, guaranteed to outperform any individual contribution? Clearly, this is only possible if the sources compensate for the errors and limitations of one another.\(^2\) Suppose all four sources report the possibility of snow (i.e., \(x = \text{“it is going snow”}\)), what would be the confidence in the resulting prediction by the agent? We now consider two different reporting settings, and their significance on the agent’s task. Before that, we assume reports from the sources are represented as opinions using subjective logic. Each report received is, therefore, of the form \(\omega^y_x = (1.0, 0.0, 0.0)\), implying that source y has absolute certainty about

\(^1\)http://www.wunderground.com

\(^2\)This is the general idea behind multi-source fusion.
the possibility of snow. Similarly, the decision-maker’s opinion about the reliability of each of the sources is represented as $\alpha^y_A = (0.8, 0.2, 0.0)$, implying that the agent $A$ has a 0.8 belief in source $y$ being reliable, and a 0.2 belief in the source being unreliable.

- **Setting I (Independent Reporting):** In this setting, it is assumed that all four sources report independently. For instance, the sources may be using different models to determine the likelihood of snow. A fusion operation is carried out, taking the reliability of each source into account to arrive at a conclusion, $\alpha^x_A = (0.94, 0.0, 0.06)$. This estimate (from $A$’s point of view), leads to a lower uncertainty, as opposed to the report of any of the sources taken individually. An important point here is, given that each source reports independently, the decision-maker’s confidence in the combined estimates is realistic, and can therefore be trusted. This argument is based on the intuition that, when there is uncertainty regarding reports from individual sources, the degree of corroboration (i.e., how many sources provide the same report) could provide a realistic indication of the trustworthiness of information.

- **Setting II (Dependent Reporting):** If the sources rely on each other’s reports, then they will fail or succeed identically. This correlated bias implies that the resulting estimate $\alpha^y_A = (0.94, 0.0, 0.06)$, will be misleading to the decision-maker. Not only that, there is the potential impact on the resources used in sampling all the sources. If, for instance, the decision-maker is able to account for the dependency among the sources, the uncertainty in the fused report should remain at $(0.8, 0.0, 0.2)$, no matter how many contributions are integrated. Moreover, sampling all the sources would be redundant, if not wasteful (in terms of resources).

Example 2 presents a simple illustration of the significance of integrating evidence from diverse sources. The reason most truth discovery approaches fall short in this context is that they assume that sampling sources is cost-free. Not only that, some of these approaches also assume that reports from the sources are provided independently (Thomopoulos et al., 1987; Teacy et al., 2008). However, many familiar social information systems can be characterised by some form of correlated behaviour among information sources. Accounting for dependencies in the truth discovery process is advantageous, not only as a means for minimising the cost of information acquisition, but also for making better assessments.

We are interested on how source diversification processes can be modelled and applied in the contexts of source selection and fusion for truth discovery. Broadly speaking, our view of diversity is a stratification of the source population, such that sources likely to provide similar reports are grouped together. We note here that our requirement for diversity goes beyond simply accounting for dependencies among information sources. We seek to capture richer information contexts, such as differences in expertise and perspectives, which may be exploited by a decision-maker to make as accurate an estimate as possible. For example, the cost and risk analysis of interacting with certain groups of sources may serve to inform how sampling decisions are made (Jøsang and Presti, 2004). Such groups may, for example, be communities in a geographic region, divisions in a corporation, or sensors owned by a specific organisation.

---

3This value was arrived at using SL’s discount and fusion operators defined in Section 3.1.5.
Thinking about diversity in populations of information or opinion providers is not a new idea; this is a common tactic used in the social sciences and by polling organisations. Shiller (1995) suggests that people who interact with each other regularly tend to think and behave similarly, and describes how, for example, political beliefs or opinions on policy issues tend to show geographical and social patterns. This is often referred to as herd mentality (or herding) (Surowiecki and Silverman, 2007; Raafat et al., 2009): the alignment of thoughts or behaviours of individuals in a group through local interactions. For example, individuals from the same organisation tend to behave in a similar manner based on certain codes of conduct or policies. People who subscribe to the same news channels tend to maintain similar views about their environment. The belief in certain countries is that some diseases are caused by microorganisms known as germs, whereas people in some other cultures believe that diseases are caused by malevolent spirits. In general, entities in different populations may have diverse beliefs about the state of the world. These populations, or subgroups, are often defined by a range of features (age, nationality, geography, etc.) that may influence their behaviour. Exploiting correlations between behaviour and observable features of agents has also been explored in computational models of trust, where the problem addressed is to whom should a task be delegated. Liu et al. (2009) use clustering techniques to learn stereotypes on the basis of past transactions and assess agents according to those stereotypes. Burnett et al. (2010) use model tree learning to form stereotypes that are used as a prior to a Beta trust model such that direct evidence, when acquired, gradually overrides the effect of the stereotype. More recently, Şensoy et al. (2014) demonstrate the use of graph mining techniques to formulate stereotypes from structured features, such as patterns in a social network, that may be used to inform trust assessments.

Our premise is that by analysing the relationships among information sources, useful metrics can be identified for discriminating between diverse behaviour patterns. In this chapter, we employ machine learning techniques to assess the similarity of sources given histories of reports from those sources, and trust-based heuristics for sampling.

4.1 The TIDY Framework

The Trust in Information through DiversitY (TIDY) framework for multi-source integration is centred around the idea of a Diversity Structure. The framework, illustrated in Figure 4.1, uses histories of reports from information sources that exhibit certain features to learn a similarity metric. This metric is then used to cluster sources on the basis of their features to form a diversity structure. A sampling strategy is then employed that is informed by this diversity structure, and reports acquired from the sampling process are fused to provide an estimate of the environmental state.

In formalising the TIDY framework, we assume a decision-maker (or agent) that has the task of monitoring an environmental state (e.g., the weather condition at a location, or the number of casualties following a disaster).

**Definition 14 (Task)** A task is the activity of deriving an estimate of some environmental state $\theta^t$, at each time $t \in T$ within an interval $[t_1, t_2]$, such that $t_2 \geq t \geq t_1$. 
4.1. The TIDY Framework

The domain of the variable $\theta$ may be different for different query types, such as “is it snowing?” and “how many casualties?”. For a particular query, $\Theta$ represents the set of possible values of $\theta$.

To acquire an estimate of $\theta$, $\hat{\theta} \in \Theta$, sources of varying trustworthiness may be queried, the result of which will be a set of reports from the selected sources.

**Definition 15 (Information Source)** An information source is a tuple $\langle x, V_x \rangle$, where $x \in \mathcal{N}$ such that $\mathcal{N} = \{1, \ldots, n\}$ is a unique identifier and $V_x$ is a vector containing values for $x$’s features.

**Definition 16 (Report)** A report, $o_x$, received from source $x \in \mathcal{N}$ is a tuple containing the measured value, $o \in \Theta$ and a confidence measure, $\delta$: $o_x = \langle o, \delta \rangle$. The set of all reports from source $x$ is $O_x$.

The confidence measure, $\delta$, is meta-information expressing the level of confidence (or uncertainty) the reporting source attaches to its report. A sensor, for example, may report the water level of a river to a specific accuracy. The decision-maker maintains histories of reports received from each information source.

**Definition 17 (Report History)** A history of reports from a source is a sequence, defined as a function $h_x : T \rightarrow O_{x \perp}$ where $O_{x \perp} = O_x \cup \{\perp\}$. If, for some $t$, $h_x(t) = \perp$, then no report was received from source $x$ at time $t$. For convenience, we refer to $h_x(t)$ as $h_x^t$, and we define the reports received at time $t$ as $O^t = \bigcup_{x \in \mathcal{N}} h_x^t$.

We assume a finite set of features that can be used to describe information sources. Examples may include the location of a source, or its organisational affiliation.
Definition 18 (Feature) Let \( F = \{ f_1, \ldots, f_d \} \) be the set of all features. A feature \( f_i \in F \) is an observable attribute of an information source.

Each feature \( f_i \in F \) has some domain \( D_i \), and for each source \( x \in \mathcal{N} \), there exists a feature value \( v_i \in V_x \), such that \( v_i \in D_i \cup \{ \text{null} \} \). If a feature is unobserved or not relevant, its value is \( \text{null} \) for that source.

In order to group sources according to their features, we need a good similarity metric that allows the decision-maker to estimate the degree of similarity between sources.

Definition 19 (Similarity Metric) A similarity metric is a function \( \text{sim} : \mathcal{N} \times \mathcal{N} \rightarrow \mathbb{R} \).

The idea behind this definition of a similarity metric for the TIDY framework is that similarity in reporting patterns may correlate with similarity in some of the sources’ features. If this is the case, then given two sources \( \langle x, V_x \rangle \) and \( \langle y, V_y \rangle \), the function \( \text{sim}(x, y) \) will give a score representing the degree of similarity that is expected in reports from these sources. This allows us to overcome the challenge of insufficient or sparse evidence regarding sources’ reporting patterns, and to easily generalise to unseen cases (Burnett et al., 2010).

We can then use this similarity metric to stratify, or cluster sources to form a diversity structure.

Definition 20 (Diversity Structure) A diversity structure, \( \mathcal{DS} \), is a stratification of the set of all sources in the system into exhaustive and disjoint groups. \( \mathcal{DS} = \{ G_1, \ldots, G_K \} \), such that \( \bigcup_{k=1}^{K} G_k = \mathcal{N} \) and \( G_k \cap G_l = \emptyset \) for any \( k, l \in \{1, \ldots, K\} \) with \( k \neq l \).

In forming a diversity structure, we assume there is some function \( \Delta \), such that, given some set of sources and a similarity metric, will compute a diversity structure; i.e., \( \mathcal{DS} = \Delta(\text{sim}, \mathcal{N}) \). This function may be realised through an off-the-shelf clustering algorithm such as hierarchical or k-means clustering (Jain et al., 1999).

4.1.1 Source Agreement

Our aim is to generalise from similarity in sequences of reports from different sources to similarity of sources on the basis of their observable features. In general, we require a function that, given histories of reports from two sources, provides an assessment of the level of agreement between the reports received from those sources. In the TIDY framework, we make the assumption that agreement between histories of reports from two sources can be derived from assessments of agreement between individual reports when reports are received from the two sources at the same time. The rationale for this is that we may then define a mechanism for assessing agreement between sources that operates efficiently on streams of reports received.

Definition 21 (Report Agreement) An assessment of the extent to which reports from two sources agree is determined by the function \( \nu_{\text{agr}} : T \times O \times O \rightarrow \Pi_{\text{agr}} \), where \( \Pi_{\text{agr}} \bot = \Pi_{\text{agr}} \cup \bot \). We require that \( \nu_{\text{agr}}(t, h'_x, h'_y) = \bot \) if either \( h'_x = \bot \) or \( h'_y = \bot \); i.e., agreement can only be assessed if we receive reports from two sources at the same time.

The report agreement function will depend on the underlying model of report and source agreement. If, for example, a Beta distribution (Jøsang, 2013) is used to model agreement,
\[ \Pi_{agr} = \{0, 1\}, \] where 0 indicates that reports do not agree and 1 that they do agree. Now, given an assessment of the extent to which two reports agree, we may define a means to compute the agreement between sources.

**Definition 22 (Source Agreement)** Agreement between sources is some aggregation of a sequence of agreements between reports that have been received from the two sources at the same time. \( \sigma : (\mathcal{N} \times \mathcal{N}) \times (T \times O \times O \rightarrow \Pi_{agr}) \rightarrow \mathbb{R} \)

With assessments of how sources agree, we may revise the function used to assess similarity between sources (Definition 19).

**Definition 23 (Similarity Metric Revision)** The similarity assessment function is revised on the basis of (dis)agreements between reports received from sources such that \( \text{revise}_\text{metric} : ((\mathcal{N} \times \mathcal{N}) \times (T \times O \times O \rightarrow \Pi_{agr}) \rightarrow (\mathcal{N} \times \mathcal{N} \rightarrow \mathbb{R}) \)

The source agreement function, \( \sigma \) provides a means to compute an agreement score of each source pair. Through the identifiers of the source pairs we have the values of the observable features of each source. The \( \text{revise}_\text{metric} \) function represents the problem of computing a classifier that assigns a similarity score for two sources on the basis of the values of their features.

### 4.1.2 Trust Assessment

In addition to source similarity, an important factor in making source querying decisions is the extent to which we trust a source to provide an accurate report. We assume that the agent is able to observe ground truth, \( \theta \), but this is only available at a time after which it is useful for decision making. This observation of ground truth may, however, be used to revise our assessments of information sources, given we have a history of reports from those sources.

**Definition 24 (Report Assessment)** The assessment of a report against ground truth is determined by the function \( \nu_{\text{tru}} : T \times O \times \Theta \rightarrow \Pi_{\text{tru}, \perp} \), where \( \Pi_{\text{tru}, \perp} = \Pi_{\text{tru}} \cup \perp \). We require that \( \nu_{\text{tru}}(t, h^\perp, \theta^t) = \perp \) if \( h^\perp = \perp \).

Again, this function will depend on the underlying model, and as with source agreement we can define a source trust assessment function.

**Definition 25 (Source Trustworthiness)** The trustworthiness of a source is determined by assessments of the sequence of reports received from that source over time, and is determined by the mapping \( \tau : \mathcal{N} \times (T \times O \times \Theta \rightarrow \Pi_{\text{tru}}) \rightarrow \mathbb{R} \).

Information about the trustworthiness of sources is recorded for each source using an appropriate instantiation of this function (e.g. a Beta probability density function).

### 4.1.3 Sampling

While monitoring the environmental state, the decision-maker will acquire reports from various sources over time. The decision-maker must make a decision on how to sample for evidence. In particular, the agent must decide which sources to sample. The objective of a sampling strategy is to select a subset, \( \mathcal{N} \subseteq \mathcal{N} \), in order to maximise its utility. The utility of a sampling decision is a function of the accuracy of the estimate \( \hat{\theta} \) (or information quality), and the cost of sampling.

The quality of information measures the degree of accuracy of an estimate of the environmental state with respect to ground truth.
Definition 26 (Information Quality) The information quality obtained from sampling a set of sources is a function: \( \text{qual} : \Theta \times \hat{\Theta} \rightarrow \mathbb{R} \). For example, if \( \Theta \) is the environmental state and \( \hat{\Theta} \) is the estimate of that state obtained, the information quality is \( \text{qual}(\Theta, \hat{\Theta}) \).

In sampling sources, the decision-maker incurs cost. We assume that the cost of sampling a specific source remains fairly stable over time, or changes at a very slow and predictable rate.\(^4\) Nevertheless, costs may vary across sources; e.g., the cost of asking an expert may be different from polling a group of friends.

Definition 27 (Sampling Cost) Sampling cost is a function: \( \text{cost} : 2^N \rightarrow \mathbb{R} \). In many settings, sampling costs are strictly additive: \( \text{cost}(N) = \sum_{x \in N} \text{cost}({x}) \).

The decision-maker’s task is to select a subset of sources in order to maximise its utility, or, more generally, maximise its expected utility over a sequence of sampling decisions given that the act of sampling provides information about the characteristics of the information sources sampled.

Definition 28 (Sampling Utility) Sampling utility is a function: \( u : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \).

4.1.4 Fusion

Fusion provides the decision-maker the necessary tools for combining reports from various sources to arrive at an estimate.

Definition 29 (Fusion) Information fusion is a function, \( \mathcal{F} : 2^O \rightarrow \Theta \), that computes an estimate of the environmental state, \( \hat{\Theta} \), given a set of reports, \( O' \).

In performing fusion, the decision-maker may take into account estimates of the trustworthiness of sources, such as in Subjective Logic (Jøsang, 2013) where these assessments are used to discount reports received from sources.

We build upon this framework for source selection and fusion in this chapter and the next. In what follows, we present a specific realisation of the framework that we evaluate in Section 4.3.

4.2 A Realisation: TIDY\(_0\)

In Chapter 3, we provided background on the techniques that form the basis of our concrete framework. TIDY\(_0\), our specific realisation of TIDY in this chapter, builds on some of these techniques. An agent can exploit them to generalise over past interactions involving the sources in order to make inferences regarding their behaviour in future interactions. For instance, if certain groups of sources have been observed to consistently agree on the same sets of issue or query, then we can generalise to unseen cases by disambiguating the salient features defining such patterns. This is important, as it will often be impracticable to interact with all sources in large, dynamic systems. In realising the framework, we revisit some of the concepts defined in Section 4.1, albeit in a less formal manner.

\(^4\)This is consistent with most real-world economic settings (Feldstein, 1997).
4.2. A Realisation: TIDY

4.2.1 Task
Generally speaking, a task refers to the process of determining what is true in the world such as the current river level. A task may be repetitive, requiring a periodic assessment of the ground situation (e.g., hourly or daily weather updates). Repetitive task models arise naturally in areas such as time series analysis (Osborne et al., 2008). As an example, we assume a weather station tasked with the provision of weather information. In order to achieve this objective, the decision-maker needs to constantly sample available sources at the location of interest for weather reports. The rationale for this is that the decision-maker is then able to form an opinion over time regarding the behaviour of the sources. This is as opposed to a one-shot task, where an agent may need only carry out a single transaction (e.g., buying a life insurance from a broker) without the need for long-term monitoring. In our experiments, we assume that interactions are ordered, and refer to each time period that an interaction occurs (i.e., querying a group of sources and deriving an estimate) as a sampling round.

4.2.2 Information Source
As mentioned, variously trusted sources may be queried about the state of the world. These sources can be soft (e.g., human) or hard (e.g., wireless sensors). They can be structured (e.g., databases) or unstructured (e.g., open-source data on the Internet). Sources may have different reporting capabilities (or expertise) depending on the context. For example, a UAV (unmanned aerial vehicle) may do a better job than a human in providing surveillance coverage of a disaster region. On the other hand, human sources may be better in differentiating between different kinds of wildlife affected in the aftermath of the disaster. Sources may also exhibit different behaviours based on a variety of reasons. For example, a sensor whose battery life is low may provide imprecise measurements, or drop packets. Also, some sources may obfuscate their reports before sharing in order to avoid revealing sensitive information, or may maliciously report misleading information in order to bias the truth discovery process.

4.2.3 Report
Reports obtained from sources are used to derive an estimate of the environmental state. A report can assume values from a wide range of domains such as binary, continuous, etc. For example, the report $o_x = 0$, is interpreted differently for the queries “is it snowing?” and “how many casualties?”. The first query belongs to a binary domain, s.t. $\theta \in \{1, 0\}$, and the reported value is taken to represent a negative opinion from source $x$ about the event snow. The domain of the second is the set of natural numbers, i.e., $\theta \in \mathbb{N}$, and the report is interpreted as no casualties. The confidence measure, $\delta \in [0, 1]$, represents a degree of confidence in the measured value, such that a 0 would indicate an absolute lack of confidence or uncertainty, and 1 indicates an absolute confidence attached to the measured value. In our evaluation, we assume that reports are continuous s.t. $\theta \in \mathbb{R}$.

4.2.4 Feature
Sources’ features represent attributes such as organisational affiliation, location, age, etc. Similar to reports, features can assume a wide range of values in both quantitative (i.e., continuous, discrete, and interval) or qualitative (i.e., nominal, ordinal) domains. The following illustration shows an example feature representation for three sources $x, y, z \in \mathcal{N}$:
4.2. A Realisation: TIDY₀

\[ F = \{\text{organisation, cost, age}\} \]

\[ V_α = \{\text{UOA, 0.11, 12}; V_β = \{\text{UOA, 0.12, null}; V_τ = \{\text{UOE, 0.6, 12}\} \]

4.2.5 Computing Source Agreement and Trust

A decision-maker can form opinions based on evidence obtained by interacting, and subsequently evaluating the behaviour of sources in the system. Evidence for computing both the agreement and trustworthiness of sources is gathered from different interaction contexts.

Evidence of Source Agreement

Evidence of agreement between pairs of sources can be obtained following a sampling activity. That is, after obtaining reports from sampled sources, the decision-maker is able to evaluate these experiences and thus update evidence parameters \( (r_{x,y}, s_{x,y}) \) (see Section 3.1.3) of the agreement for each source pair, \( x, y \). Using Equation 4.1, both the positive \( (r_{x,y}) \) and negative \( (s_{x,y}) \) evidence parameters can be updated in light of new evidence obtained at time step \( t \). The parameter, \( δ_{agr} \) represents an application-specific threshold for the agreement between two reports.

\[
( r'_{x,y}, s'_{x,y} ) = v_{agr}(t; h'_x, h'_y) = \begin{cases} 
(1,0), & \text{if } |h'_x - h'_y| \leq δ_{agr} \\
(0,1), & \text{if } |h'_x - h'_y| > δ_{agr} \\
(0,0), & \text{otherwise}
\end{cases} \tag{4.1}
\]

Evidence of Source Trustworthiness

After an estimate of the environmental state has been made, we assume that the decision-maker is able to observe ground truth, \( θ' \). The reliability of a source can be assessed on the basis of the conformity of its report to fact. Evidence used in computing the trustworthiness of a source, \( x \) is accumulated over time as a \( r_{x,t} \) pair (again, see Section 3.1.3). Each experience is obtained using Equation 4.2, where \( δ_{tru} \) is an application-specific threshold value for report reliability.

\[
( r'_{x,t}, s'_{x,t} ) = v_{tru}(t; h'_{x,t}, θ') = \begin{cases} 
(1,0), & \text{if } |h'_{x,t} - θ'| \leq δ_{tru} \\
(0,1), & \text{if } |h'_{x,t} - θ'| > δ_{tru} \\
(0,0), & \text{otherwise}
\end{cases} \tag{4.2}
\]

Having described how evidence may be aggregated, we now describe how these experiences may be used by an agent to compute both source agreement and trust.

The Beta distribution (Jøsang, 2013) provides a means of forming opinions based on available evidence. For instance, opinions about the degree of agreement of two sources, \( x \) and \( y \) can be formed on the basis of positive \( (r_{x,y}) \) and negative \( (s_{x,y}) \) evidence. These opinions may be updated in light of new evidence. The pair \( \langle r_{x,y}, s_{x,y} \rangle \), provides a source of \( α_{x,y} \) and \( β_{x,y} \) parameters of the Beta distribution such that: \( α_{x,y} = r_{x,y} + 1 \) and \( β_{x,y} = s_{x,y} + 1 \). The expected value of Beta(\( σ_{x,y} | α_{x,y}, β_{x,y} \)) can be derived using these parameters:

\[
E(σ_{x,y}) = \frac{α_{x,y}}{(α_{x,y} + β_{x,y})} \tag{4.3}
\]
Similarly, opinions about the trustworthiness of a source, \( x \) can be formed on the basis of positive (\( r_x \)) and negative (\( s_x \)) evidence, which may also be updated as new evidence becomes available. The expected value of Beta(\( \tau_x \mid \alpha_x, \beta_x \)) can be derived:

\[
E(\tau_x) = \frac{\alpha_x}{\alpha_x + \beta_x}
\]  

(4.4)

If considering the trustworthiness of a group of sources, \( G_i \), then group trust \( \tau_i \) can be calculated as the average trust score of group members:

\[
\tau_i = \sum_{x \in G_i} \tau_x
\]

(4.5)

While the agreement assessment we have described above provides evidence of similarity among known sources, we still require mechanisms for generalising from this evidence to a structure that can be used to stratify sources based on their observable features.

### 4.2.6 Learning a Similarity Metric

One of the main problems of relying on features of sources as useful indicator of similarity, is that an agent cannot know in advance the relative importance of different features for capturing the diversity in the population. For instance, while geographical location might be readily identified as an informative feature for stratification in political settings, in other contexts, relationships among sources might be captured by far more complex features or feature combinations not easily identified by an agent. For example, the different subgroups captured in Figure 4.2, can be described using a combination of feature relationships. “Group A” for instance, points to the importance of the composite feature \( \text{organisation} \wedge \text{age} \wedge \text{location} \) in defining similarity, whereas the single feature “age” or the composite feature \( \text{age} \wedge \text{location} \) are uninformative, and therefore not useful metrics for stratification. While these kinds of relationships may be easy to identify in some settings, in general we cannot presume that this sort of knowledge is available to a decision-maker. Instead, we focus on learning these relationships using appropriate machine learning techniques.

Decision trees (Breiman et al., 1984) provide an appropriate representational abstraction for
modelling a similarity metric. They are classification tools, which allow a label to be found for a given input by selecting paths through a tree based on conditions specified at branching nodes. Each node of a decision tree represents a particular feature, and branches from nodes are followed depending on the value of the feature represented by that node. In our own case, each input feature value holds the *distance* between the values of that feature for a source pair. This intuitively captures the notion of similarity in features. Each leaf of the tree represents a similarity score (or a function producing a similarity score), which is assigned to every source pair or classification examples reaching that leaf.

Classical decision tree induction techniques are not suitable for problems where the class value to be predicted is real-valued (Frank et al., 1998; Quinlan, 1992). As our aim is to estimate the degree of similarity between sources represented by a real-valued similarity score, we require a decision tree induction technique which accommodates real-valued class labels. One possible technique that can be employed for this is model tree learning, which allows us to learn a classifier capable of predicting similarity scores from a real-valued domain (Witten and Frank, 2005; Quinlan, 1992).

In model tree learning (see Section 3.2.3), the leaves of a tree are linear regression models, which can be used to estimate a target value. Using this technique, a similarity metric can be induced by using training examples from features of sources as well as available evidence from their report histories. We make use of the M5 model tree learning algorithm (Witten and Frank, 2005; Quinlan, 1992).

In table 4.1, we present an example of 10 training instances from which a similarity metric may be induced. Each training instance to the M5 algorithm is of the form

\[ \langle \text{dis}(v_{1x}, v_{1y}), \ldots, \text{dis}(v_{dx}, v_{dy}), \sigma_{xy} \rangle, \]

representing the feature value distances of a source pair and their degree of agreement, \( \sigma_{xy} \), as the class label. For each feature \( f_i \in F \), we obtain a value for a source pair \( x, y \) as \( \text{dis}(v_{ix}, v_{iy}) \), where \( v_{ix} \) is the value of feature \( f_i \) for source \( x \). Any suitable distance function (e.g., Euclidean distance) can be employed for this task. Our specific instantiation computes the distance for each feature value \( \text{dis}(v_{ix}, v_{iy}) \) for a source pair, \( x, y \), as the absolute

---

We use the M5 implementation of Weka (Hall et al., 2009), a popular open-source machine learning toolkit written in Java.

---

### Table 4.1: Training examples

<table>
<thead>
<tr>
<th>#</th>
<th>organisation</th>
<th>cost</th>
<th>age</th>
<th>power</th>
<th>( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>0.11</td>
<td>2</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>0.6</td>
<td>3</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>0.12</td>
<td>1</td>
<td>null</td>
<td>0.7</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>0.99</td>
<td>1</td>
<td>0.9</td>
<td>0.3</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0.11</td>
<td>4</td>
<td>null</td>
<td>0.89</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>0.79</td>
<td>4</td>
<td>0.54</td>
<td>0.4</td>
</tr>
<tr>
<td>7</td>
<td>20</td>
<td>0.01</td>
<td>2</td>
<td>0.2</td>
<td>0.78</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>0.89</td>
<td>5</td>
<td>0.6</td>
<td>0.45</td>
</tr>
<tr>
<td>9</td>
<td>100</td>
<td>0.33</td>
<td>2</td>
<td>0.7</td>
<td>0.5</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>0.1</td>
<td>5</td>
<td>0.24</td>
<td>0.8</td>
</tr>
</tbody>
</table>
difference of their features; \( dis(v_{i,x}, v_{i,y}) = |v_{i,x} - v_{i,y}| \). Where a feature has a null value for either or both sources, no reasonable comparison can be made. In such an instance, the entry for the feature value for the source pair is assigned a null value. The agreement between a pair, \( \sigma_{x,y} \), is derived using Equation 4.3. This value is 0.5 when the system is initially instantiated, denoting an equal likelihood of both outcomes (i.e., agreement and disagreement) before any positive (agreement) or negative (disagreement) evidence is observed \( (r_{x,y} = s_{x,y} = 0) \). Lack of evidence may likely impact the ability of the learned metric in identifying the desired correlations. Therefore, it is necessary that a revision of the model be carried out periodically as evidence is accumulated through repeated interactions.

As well as the ability to handle features of different kinds (e.g., nominal), the model tree algorithm is robust to missing values that pose a risk of overfitting a learned model (Witten and Frank, 2005, p. 86). This is particularly important to us, given that some of the sources might not have values for certain features in \( F \). For instance, a human source might not have a value for the feature power, which may be relevant to other kinds of sources (e.g., wireless sensors). One way of handling this challenge is to use the class value as a surrogate attribute in a training set. In a test set, one possible solution is to replace the unknown attribute value with the average value of that attribute for the training examples that reach the node.

An example similarity metric induced using the training set in Table 4.1 is shown in Figure 4.3. The linear models at the leaf nodes are linear combinations of the attributes with assigned weights, and are of the form: \( w_0 + w_1 f_1 + w_2 f_2 + w_3 f_3 \), where \( w_j \) is a scalar weight, and \( f_j \) is a feature. This metric can be used to classify pairs of sources by tracing a path through the tree in order to determine an appropriate linear model to employ. The output is a real value that represents the similarity score for the source pair. Using this structure, an agent can easily generalise to a

---

6The M5 algorithm can accommodate other feature types including qualitative. Also, different metrics exist for computing the distance between features of other kinds (Ichino and Yaguchi, 1994).
notion of similarity on the basis of sources’ features, and thus being equipped with a useful tool to form a diversity structure.

4.2.7 Creating a Diversity Structure

To form a diversity structure, \( DS \), we employ hierarchical clustering (Jain et al., 1999). This is a well-known technique that can be employed for group formation. In contrast with other clustering techniques such as \( k \)-means clustering, hierarchical clustering allows us to cluster into a set of groups the cardinality of which we do not know in advance. The stratification uses an agglomerative method (see Section 3.3.4), where each source starts in a singleton group. The proximity between each source pair, \( x, y \) is then computed using the similarity metric, \( \text{sim}(x, y) \). That is, given their feature vectors, an appropriate regression model can be selected for obtaining a similarity score. The two most similar groups are then merged. Merging of groups continues until a stopping criterion is satisfied. One can, for instance, decide to stop either when the groups are too far apart to be merged (distance criterion) or when there is a sufficiently small number of clusters (number criterion). We model the stoppage criterion using a (diversity) threshold parameter, \( \psi \). This parameter value lies in the interval \([0, 1]\), and specifies the maximum level of diversity required in the system. For instance, if \( \psi = 1 \), all the sources will be assigned to singleton groups; a condition of extreme diversity. On the other hand, if \( \psi = 0 \), all the sources are assigned to one group; a condition of no diversity.

It is important to note that different similarity metrics may lead to different diversity structures, and the number of possible stratifications increases exponentially with the number of sources. For example, in a population with four sources, there are 15 possible groups (i.e., \( 2^N - 1 \), \( N = 4 \) sources), and 15 possible diversity structures as illustrated in Figure 4.4. In this research,
our focus is not on finding an optimal diversity structure. This problem has received wide attention in areas such as coalition formation (Farinelli et al., 2013), where the primary objective is to calculate and optimise the value of a coalition (group).

4.2.8 Model Validity

A diversity structure once constructed, provides a static estimate of appropriate grouping of sources in the population. Sources and their availability may, however, change. New sources may appear and sources may become unavailable. Although we can assign new sources to groups on the basis of their features rather than waiting for behavioural evidence, this does require us to consider which cluster is the best fit for any new source. Unavailable sources can simply be removed from their clusters. The behaviour of sources may also change over time, which may warrant a revision to the model of their relative similarity. New evidence from the behaviour of sources in previously unseen situations may also provide evidence that could lead to a more refined similarity metric.

One way of incorporating fresh evidence would be to revise the model periodically by defining a learning interval $L$. This interval may be determined by the number of interactions the decision-maker carries out with the environment before invoking the $\text{revise\_metric}$ function.

In revising the model, new examples, $\langle \text{dis}(v_{1,x}, v_{1,y}), \ldots, \text{dis}(v_{d,x}, v_{d,y}), \sigma_{x,y} \rangle$ are added to the training set for each source pair in the population, and the model tree is then reconstructed. It is not necessarily the case that features of sources would change over time. For example, while it is possible that features such as battery-life (of say, wireless sensors) may change over time, other features such as ownership may remain fairly stable over a period of time. Evidence of source agreement accumulated over time as a $\langle r_{x,y}, s_{x,y} \rangle$ pair is used to obtain an updated agreement score (see Equation 4.3).

Although quite straightforward, employing a learning interval for model revision is insensitive to the dynamics in the population, and therefore may lead to unnecessary overheads. It is preferrable that the diversity model be revised based only on evidence. For instance, available evidence may suggest merging groups previously thought to be different, given the high rate of agreement in the reports of sources belonging to those groups. There may also be evidence suggesting the need to split certain groups in which members are observed to disagree a lot. An agent could, for example, employ a threshold level of error that a current model should operate within before being revised. These sort of evidence-based revisions are necessarily heuristics, and have the advantage of being much quicker than rebuilding the model from scratch. However, a limitation with this approach is the chance of anomalous revisions which may not adapt well to the global population.

We have discussed one possible reaction of an agent to changes in the source population. However, that does not preclude other forms of responses from a decision-maker. For instance, instead of always resorting to model revision, the decision-maker may adapt its sampling strategies in line with evidence pointing to the current state of the model. A decision-maker may, for example, decide to sample more from groups where available evidence suggests higher rates of disagreement in the reports and vice versa. This is a non-trivial decision problem (Etoré and Jourdain, 2010; Zheng and Padmanabhan, 2002), which we seek to address in Chapter 5. However, for the purpose of our instantiation in TIDY$_0$, the use of a learning interval is adopted.
4.2.9 Sampling
The primary objective of a diversity structure is to aid the process of source selection. The use of a diversity structure as a basis for sampling has some similarities to stratified sampling (Cochran, 1977). Stratified sampling has been shown to perform well in many survey applications including social media analytics (O’Connor et al., 2010). It involves partitioning a population into disjoint subgroups according to some stratification variable (e.g., geography, culture, age). Samples are then taken from each subgroup independently to estimate the variable of interest. While similar in some aspects, the sampling strategy we propose in this research is significantly different from stratified sampling.

The number of groups in a diversity structure is $|DS|$. For each $G_i \in DS$, the subset of sources sampled is $g_i \subseteq G_i$. The set $\mathcal{G}$ contains all the groups sampled. We consider two sampling strategies, contingent on a sampling budget, $\Phi$ (defined in Section 4.3.1):

- **Strategy I ($\Phi \geq |DS|$)**: The number of candidates to be sampled, or the budget assigned to a group $G_i$ is determined by the size of the group:

  \[ \text{budget}(G_i) = |G_i| \times \left( \frac{\Phi}{|\mathcal{N}|} \right) \]  

  Individual sources are then randomly selected from $G_i$ according to this budget. Applying this technique may, however, lead to information exclusion in much smaller groups (e.g., not selecting from singleton groups). For this reason, we select at least one candidate from each group, and correspondingly reducing the number of candidates to be sampled from much larger groups.

- **Strategy II ($\Phi < |DS|$)**: This strategy is applied only if the sampling budget is insufficient to cover all groups. Groups are ranked in order of members’ trustworthiness (using Equation 4.5). Then a single source is selected from the most trustworthy group, then the second most trustworthy group and so on until the budget is exhausted. The intuition here is that, although information is lost from some of the groups, it is more beneficial for a decision-maker to prioritise available resources to more trustworthy groups.

We do not suggest these to be the only methods for sampling. We have selected these heuristic methods because they exploit our source diversification mechanism, and allow us to assess (through experiments) the merits of learning a diversity structure. In addition, these sampling methods are reasonable heuristics that are related to practical survey methods.

4.2.10 Fusion
Reports from sampled sources are combined in order to derive an estimate, $\hat{\theta}$. The reports from sources within a group, $G_i$, are aggregated to form a group estimate, $\hat{q}_i$:

\[ \hat{q}_i = \sum_{x \in G_i} \frac{o_x}{|G_i|} \]  

The resulting estimates from each of the groups sampled are then discounted by their corresponding trust scores, $\tau_i$. Finally, the normalised opinions from all groups are combined to obtain the estimate, $\hat{\theta}$:
4.3 Evaluation

We are interested in understanding the effectiveness of a diversity-based sampling approach to truth discovery in resource-constrained environments where information sources vary in trustworthiness. To explore this, we conducted two sets of experiments: in the first set, the independent variables are sampling budget and the proportion of malicious sources (i.e. sources that are more likely to provide misleading but independent reports); in the second, the independent variables are sampling budget and the proportion of colluding sources (i.e., sources that are more likely to copy each other’s reports). In each case the dependent variable is the mean absolute error in the resulting estimate of the environmental state.

The following hypotheses form the basis of our evaluation:

- **Hypothesis 1**: With a higher percentage of malicious sources under varying budgetary constraints, an agent using a diversity-based approach is able to make better assessments of ground truth than those that do not.

- **Hypothesis 2**: With varying degrees of source dependence and budgetary constraints, an agent using a diversity-based approach is able to make better assessments of ground truth than those that do not.

To test these hypotheses, we compare the following models of truth discovery:
4.3. Evaluation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>100</td>
<td>No. of sources in popl.</td>
</tr>
<tr>
<td>$P_l$</td>
<td>0.1</td>
<td>Popl. change probability</td>
</tr>
<tr>
<td>$\psi$</td>
<td>0.4</td>
<td>Diversity threshold</td>
</tr>
<tr>
<td>$L$</td>
<td>30</td>
<td>Learning interval</td>
</tr>
<tr>
<td>$\delta_{agr}$</td>
<td>0.1</td>
<td>Report agreement threshold</td>
</tr>
<tr>
<td>$\delta_{tru}$</td>
<td>0.1</td>
<td>Report reliability threshold</td>
</tr>
</tbody>
</table>

Table 4.2: Experimental parameters

**Diversity-Based Sampling (DBS)** Diversity-based sampling uses the realisation of the TIDY framework, TIDY$_0$, defined above.

**Observation-Based Sampling (OBS)** Observation-based sampling uses assessments of the trustworthiness of individual sources to guide sampling. This is a common approach in trust-based service selection models (Ganeriwal et al., 2008; Jøsang et al., 2007; Teacy et al., 2006). Various algorithms have been proposed, but we model the trustworthiness of each source using a Beta distribution, which is exactly the same trust model used by DBS. This allows us to effectively assess the merits of our diversity model, without having to worry about managing the complexity or high computation costs associated with models such as HABIT (Teacy et al., 2012). In addition to driving source selection, trust assessments are used to discount reports received during fusion; an approach referred to as *exogenous* discounting (Jøsang et al., 2007). When constrained by budget, OBS selects the most trusted sources according to the budget allowance.

**Majority-Based Sampling (MBS)** Majority-based sampling is based on *endogenous* filtering (Jøsang et al., 2007). This technique uses the statistical properties of the reports themselves as a basis for assessment (Zhang et al., 2006; Whitby et al., 2004). In fusion, reports deviating from mainstream opinion are filtered out. In particular, MBS filters out reports that deviate more than one standard deviation from the mean report. Therefore, estimation of the environmental state is based on the mean report of the selected sources. In source selection, sources that are closer to the majority (mean) opinion are selected preferentially.

**Random Sampling (RBS)** Random-based sampling is a popular method in conducting surveys (Waksberg, 1978). In this approach, each source has an equal probability of being sampled irrespective of previous performance. Similar to MBS, RBS estimates the environmental state using the average report of the sampled sources. The difference being that it does not perform filtering (as in MBS) or weighting (as in OBS).

4.3.1 Experimental Environment

A summary of the experimental parameters used is provided in Table 4.2. Each information source in our experiments is assigned a profile, which determines its reporting pattern in relation to other sources in the system. Each profile has three features, and for each feature, a distribution is defined from which feature values may be drawn for individual sources in the profile. Each feature value
is drawn from a Gaussian distribution, with informative profile features distributed according to $N(\mu, 0.01)$, and uninformative profile features distributed according to $N(\mu, 1.0)$. In addition, each profile has a conformity parameter, $P_c$, that specifies the degree to which reports of sources in a profile tend to be correlated. Therefore, with probability, $P_c$, a source will provide a similar report to other profile members, and with probability $1 - P_c$, it provides an independent report. Specifically, a source that does not conform, deviates from mainstream opinion held by its profile. A low $P_c$ value means that more sources in a profile will report independently, according to their individual reliability model. A conforming source when reporting, first finds out about opinions maintained by its profile members. If any exists, it randomly selects one of these opinions to report, discarding its own private opinion. In this way, we model the correlated bias among sources we wish our model to identify. The $P_c$ parameter adds an extra challenge to the learning algorithm, and allows us to evaluate the ability of our model to cope with noise due to uncorrelated feature-behaviour similarity. Unless otherwise stated, the $P_c$ parameter is set at 0.8 for all profiles. A summary of the profiles is provided in Table 4.3. In the table, informative features for defining similarity are marked with an “x”, while unmarked ones are noise features.

Since our work is also placed in the context of large and open environments, sources may freely join and leave the system at any time. We model this condition in the system using the population change probability, $P_l$. Specifically, $P_l$ is used to specify in each sampling round, the probability that a source will leave the system. When a source leaves, it is replaced with a new source of the same profile in order to keep the number of sources fixed throughout the simulation. This property impacts on the ability of the different approaches to accurately model the behaviour of sources and emphasises the need for a good exploration of the population. However, dynamic activity is relaxed in all cases for the first 30 sampling rounds of each experiment to enable the different approaches to gather information to build their individual models. By setting the $P_l$ parameter at 0.1, we provide a reasonably stable environment that enables the behaviour of the sources to be modelled. A $P_l$ value higher than 0.1 leads to a highly dynamic system, with little or no opportunity for learning and modelling source behaviour.

Each source has a reliability parameter, $P_r$ that determines the type of reports it provides (i.e., honest, malicious). We define the following report types:

- **Reliable report**: This type of report is closer to ground truth or fact, $\theta$, and is drawn from the distribution $N(\theta, 0.01)$. Reports from sources with high reliability ratio, $P_r$ are more likely to fall into this category.

<table>
<thead>
<tr>
<th>ID</th>
<th>$f_1$</th>
<th>$f_2$</th>
<th>$f_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_1$</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p_2$</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>$p_3$</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>$p_4$</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>$p_5$</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

*Table 4.3: Source profiles*
4.3. Evaluation

Figure 4.6: Increasing proportion of malicious sources with different budget (Φ) constraints
4.3. Evaluation

- Malicious report: Reports of this kind are significantly deviated from ground truth or fact, and follow the distribution $N(\theta + 1, 0.01)$. Reports from sources with low $P_r$ are more likely to fall into this category, which, if left unmanaged, could potentially undermine the truth discovery process.

The report reliability threshold, $\delta_{rel}$ is set to 0.1, which reflects the intuition that information is still useful if it has a small amount of noise or is slightly discounted (Şensoy et al., 2013). In line with this, the report agreement threshold, $\delta_{agr}$ is set to 0.1 to reflect the intuition that a source pair is considered to agree if their reports are slightly deviated from each other’s.

To permit a clear discussion and evaluation of our source diversification model, we assume a setting with a fixed budget, $\Phi$, such that: $\text{cost}_N \leq \Phi$. In particular, we define budget, $\Phi$, in terms of the number of sources that may be sampled for evidence, such that the subset $N \leq \Phi$. Consequently, we define a small budget as $\Phi = 5$, a medium budget as $\Phi = 25$, and a large budget as $\Phi = 75$. This allows us to evaluate the performance of the different approaches under different sampling constraints.

4.3.2 Results

Each instance of our simulation was repeated 10 times, with each run having 100 sampling rounds. Statistical significance of differences between strategies was computed using analysis of variance (ANOVA). Analyses of significant differences between pairs of strategies were carried out using Tukey’s HSD ($\alpha$-level = 0.05). We present and analyse the mean absolute error (information quality) averaged over multiple runs for the different strategies considered. Error bars represent a 95% confidence interval (CI) variation between means.

Hypothesis 1

Our first set of experiments is focused on assessing the ability of our model to provide good estimates in the presence of malicious sources. The null and the alternative hypotheses are:

- $H_0$: There is NO significant difference (two-tailed) in the performance of the different truth discovery models with a higher percentage of malicious sources increases under varying budgetary constraints.

- $H_1$: There is a significant difference (two-tailed) in the performance of the different truth discovery models with a higher percentage of malicious sources increases under varying budgetary constraints.

Figure 4.6 shows how our model, DBS, compares to other models under varying budgetary constraints. The small budget setting, shown in Figure 4.6(a), shows that Majority-Based Sampling (MBS) has a slight edge in performance when the proportion of malicious sources is low ($< 0.3$). This is because MBS benefits from the high proportion of honest sources, who are likely to provide reliable reports used in filtering out bogus ones. In addition, the approach is not affected by the dynamic nature of sources in the system (modelled by the $P_l$ parameter); filtering is based only on the statistical properties of the reports, not on any knowledge of source behaviour. The performance lag in the case of Diversity-Based Sampling (DBS) and Observation-Based Sampling (OBS) can be attributed to the discounting of opinions. Both approaches use the trust scores of sources as discounting weights. When the correct weights are not known, reports from sources
could be misrepresented. This problem is amplified by the dynamic nature of sources in the system, thus making it challenging for the decision-maker to determine the true reliability of sources, hence appropriate weights for their reports. This observation in itself suggests that in environments with low proportion of malicious sources, discounting may lead to poor estimates of ground truth, especially when discounting weights are not appropriately tuned.

The performance of the MBS approach is observed to degrade significantly as the percentage of malicious sources increases (starting from 0.3). This result is expected, since majority-based sampling approaches are not robust in the presence of increasing number of malicious reports. On the contrary, DBS shows a steady increase in performance under as the proportion of malicious sources starts to increase. The statistical analysis suggests that there is a highly significant difference between the performance of the different approaches ($p = 7.55 \times 10^{-6}$) when the proportion of malicious is high ($\geq 0.6$). This leads to the rejection of the null hypothesis. We conclude that there is a significant difference between the performance of the different approaches. A post hoc analysis (using Tukey’s HSD) allows us to examine specifically how our model compares to the other approaches in terms of performance difference. The test: DBS vs. OBS records an adjusted $p$-value of 0.003. This suggests a highly significant difference between the performance of both the DBS and OBS approaches, with DBS having on average an estimation accuracy of about 24% higher than OBS. The Random-Based Sampling (RBS) is equally outperformed by DBS. The adjusted $p$-value for the test DBS vs. RBS is $\ll 0.001$, with DBS having on average 39% higher estimation accuracy. The test: DBS vs. MBS suggests that there is also a highly significant performance difference between DBS and MBS. The adjusted $p$-value is $\ll 0.001$, with DBS having on average about 44% higher estimation accuracy. These results demonstrate that in contexts of limited budget, a model of diversity leads to better assessments in the presence of a high proportion of malicious sources. As observed in Figure 4.6(a), both DBS and OBS approaches tend to perform better than the other approaches under this setting. This observation points to the merits of discounting when the proportion of malicious sources is high in the system. The MBS technique on the other hand, completely falls over when the majority of sources are unreliable. This technique is observed to be out-performed even by the random selection strategy (RBS), which may, at certain times select a reliable source purely by chance. Performance of DBS remains relatively stable with increasing proportion of malicious sources. In comparison to OBS, DBS is much more robust to a dynamic population: it exploits knowledge of the groups of unknown sources to appropriately evaluate their reports.

Figure 4.6(b) shows the condition with medium budget. Again, the MBS approach performs better than the other approaches when malicious sources are in the minority ($\leq 0.3$). This result reinforces our earlier observation that modelling source behaviour isn’t necessarily advantageous when the proportion of malicious sources is low. However, DBS once again shows a consistent higher performance when there is a high proportion of malicious sources ($\geq 0.5$). The statistical analysis suggests that the performances of the different approaches have a highly significant difference ($p = 4.38 \times 10^{-6}$). A post hoc analysis indicates that DBS performs significantly better than OBS, with a 26% higher accuracy level. It also performs significantly better than RBS and MBS with 38% and 50% higher accuracy levels respectively.
Again, MBS is observed to outperform all the other approaches when the proportion of malicious source is $< 0.4$ in the large budget instance captured in Figure 4.6(c). With larger sampling budgets, filtering approaches are able to gather more evidence with which to make their assessments. In particular, MBS is able to sample more, with a higher likelihood of selecting honest sources. This enables this technique to better filter out (the few) outliers in the set. When compared to the other approaches (OBS and RBS), DBS shows no significant difference in performance under this context. However, with a high proportion of malicious sources ($\geq 0.5$), DBS again outperforms the other techniques. The statistical analysis suggests that the performance of the different approaches have a highly significant difference ($p = 5.16 \times 10^{-5}$). A post hoc analysis shows that DBS performs significantly higher than the other techniques, with DBS having on average 26%, 34%, and 49% better estimation accuracy than OBS, RBS, and MBS respectively.

In summary, we can make the following conclusions regarding hypothesis 1:

- A diversity-based approach leads to a significantly better performance (estimation accuracy) with a high proportion of malicious sources under varying budgetary constraints.

- With the exception of the majority filtering approach (MBS), the diversity model performs no worse than any of the benchmark approaches when the proportion of malicious sources is low.

**Hypothesis 2**

The degree of corroboration of evidence is often used as an indication of trustworthiness, especially in systems where there are no clear experts. In such scenarios, for example, one would be more likely to believe an event reported by numerous sources than conflicting evidence supplied by only a few sources. This is the case in different kinds of crowdsourcing applications (Kittur et al., 2008; Burke et al., 2006). If those sources are, however, simply relaying what they heard from others, then this may lead to correlated biases and misinformation.

In this set of experiments, we demonstrate the robustness of our source diversification model to varying degrees of source dependence. There are no clear experts, and the decision-maker relies on the degree of corroboration of reports to estimate environmental states. We vary the degree of source dependence from 0 to 0.9, where 0 represents a lack of dependence and 0.9 represents high dependence. The null and the alternative hypotheses are:

- $H_0$: There is NO significant difference (two-tailed) in the performance of the different truth discovery approaches under varying degrees of source dependence ($P_c$) and budgetary constraints.

- $H_1$: There is a significant difference (two-tailed) in the performance of the different truth discovery approaches under varying degrees of source dependence ($P_c$) and budgetary constraints.

We present the analysis of our results, which demonstrate the significance of our source diversification model under this setting. Figure 4.7 shows the performance of our model in comparison to the baseline approaches.
Figure 4.7: Increasing degree of source dependence with different budget (Φ) constraints
4.3. **Evaluation**

Figure 4.7(a) shows the condition with small sampling budget. When the degree of source dependence is very low (≤ 0.2), our model, DBS, tends to perform similarly to the other approaches. The only exception being when compared to MBS, which is observed to perform far worse than DBS and the other baseline approaches.

With an increasing degree of source dependence, DBS clearly outperforms all the baseline approaches. In particular, the performance of OBS is observed to be increasingly worse in comparison to that of DBS as the degree of source dependence increases. Two factors may explain this observation. First, in settings where sources are not necessarily reliable or unreliable, OBS cannot easily exploit models of sources like before to gain competitive performance. Second, since OBS approaches usually assume independence, they are not robust to correlated biases present in the source population. Although the trust component of DBS also does little in exploiting knowledge of source reliability, by modelling the diversity among sources, it can better select candidate sources for fusion in a way that is sensitive to the correlations among the sources. Also, by using *local fusion* based on identified groups, the effect of correlated biases in the final estimate is minimised.

The diversity-based approach also performs better when compared to majority-based approach, MBS. Under the small budget setting, DBS is observed to consistently outperform MBS through all degrees of dependencies. On the other hand, MBS shows a much worse performance than the other approaches. Since there are no clear experts, mainstream opinion becomes inadequate for filtering out outliers. As the proportion of dependent sources increases in the system, MBS is more inclined towards opinions held by larger groups of sources. This can be problematic under the considered setting. First, larger opinion clusters are not necessarily reliable, given the lack of experts. Second, by not aiming at diverse sources, MBS, and in fact OBS cannot effectively compensate for the errors in individual reports.

Interestingly, the RBS approach copes much better under this setting than OBS and RBS as shown in Figure 4.7(a). By randomly sampling the population RBS may, by chance, sample diverse sources thus being better able to cope with correlated biases. The statistical analysis suggests that the performances of the different approaches have a highly significant difference under the small budget setting ($p = 1.18 \times 10^{-5}$). A *post hoc* analysis indicates that DBS outperforms all the baseline approaches.

Performance of all the approaches is affected by budget. The graphs (Figure 4.7) show that performance tends to improve with an increase in budget. This suggests that a more flexible budgetary constraint enables an agent to cope much better with correlated biases. Increasing budget also impacts on the relative performance of our model to the baseline approaches. While our model does not necessarily experience a performance lag, the other approaches are better equipped, with an increased budget, to mitigate the effect of correlated biases. That notwithstanding, DBS continues to outperform all the baseline approaches under varying degrees of source dependence both in the cases of medium budget (Figure 4.7(b)) and high budget (Figure 4.7(c)). Even when sampling more sources under these budgetary conditions, the baseline approaches become vulnerable to the effect of correlated biases once the proportion of dependent sources
increases in the system. The statistical analysis suggests that the performances of the different approaches, both under the medium and high budgetary constraints, have a highly significant difference \((p = 0.0138 \text{ and } 0.0372)\) respectively. A post hoc analysis also indicates that DBS performs significantly better than the OBS, RBS, and MBS approaches.

In summary, we can make the following conclusions regarding hypothesis 2:

- With varying degrees of source dependence, a model of diversity performs significantly better in making more accurate estimates of ground truth than approaches that are not based on diversity.

- The merits of the diversity model tend to diminish with a low proportion of dependent sources. However, the model does not perform any worse than approaches that are not based on diversity.

**Sensitivity to parameter settings**

As earlier mentioned, the diversity threshold, \(\psi\), allows us to control the process of group formation. In particular, \(\psi \in [0, 1]\) models a stoppage criterion during the merging of groups to form a diversity structure. In our earlier experiments, we set this parameter at 0.4 (see Table 4.2), which, we believe, provides a reasonable cut-off mark for the identification of groups capable of accommodating different degrees of source dependence, \(P_c\). If for instance, \(\psi\) is set too high, we face the risk of not identifying (or forming) groups, even when evidence of group effect exists in the system. On the other hand, setting this parameter too low may lead to assigning all sources to a single group. This may not be the desirable outcome, especially if it does little to reflect the underlying group effect in the system. To test the sensitivity of our model to different parameter values, we conduct a separate set of experiments, with a different \(\psi\) value of 0.6. We compare the performance of our model, DBS when \(\psi = 0.4\), as used in previous experiments, and when \(\psi\) is set at 0.6. We label these instances DBS (0.4) for \(\psi = 0.4\), and DBS (0.6) for \(\psi = 0.6\) for ease of reference.

Figure 4.8 shows how the two settings DBS (0.4) and DBS (0.6) compare under varying budgetary constraints. Performance of both instances appears similar under the small budget condition (Figure 4.8(a)), with DBS (0.6) having a slight edge over DBS (0.4). However, the performance of DBS (0.4) appears to be slightly more stable with increasing proportion of malicious sources. The performance of DBS (0.4) is also observed to be more stable than that of DBS (0.6) under the medium budget condition (Figure 4.8(b)), with a similar trend being observed as well in the large budget condition in Figure 4.8(c). With large sampling budget, the performance of DBS (0.6) is observed to degrade significantly from that of DBS (0.4), when the proportion of malicious sources is high (> 0.6).

One possible explanation for the instability in the performance of DBS (0.6) as opposed to that of DBS (0.4), lies in the nature of groups identified under this setting. While DBS (0.4) is able to produce groups that more closely reflect the underlying source profiles, this isn’t the case when \(\psi\) is set at 0.6, as captured by DBS (0.6). For instance, DBS (0.6) may not be able to group similar sources together given the high diversity threshold of 0.6. In such an instance, sources that would otherwise have belonged to the same group are identified as such. This may, for example,
Figure 4.8: Comparing different $\psi$ (diversity threshold) parameter values: 0.4 and 0.6
lead to inaccurate discounting weights assigned to reports from certain groups, thereby leading to unstable performance.

4.4 Discussion

We have demonstrated that a model of diversity can lead to more accurate estimates of environmental states under varying budget constraints. This is particularly the case in contexts of dependency among information sources. It is encouraging that even when generalising on the behaviour of sources, our approach still performs as well as classical trust approaches, the focus of which is to model the behaviour of individual sources.

One drawback of the model is the sampling technique adopted. By sampling proportionally according to the size of groups, an agent may waste resources. For example, a decision-maker may allocate unnecessary budget to large but unreliable groups. The sampling strategy could be designed in a manner that takes the dynamics in the various groups into account when sampling them. We present a more rigorous approach for source selection in Chapter 5.

The source grouping mechanism we have proposed assumes that sources will exhibit the same sort of correlations. Sources may, for instance, behave differently or show different kinds of affinity in different query contexts. For example, a source may respond differently when a question concerns national interest, as opposed to one that concerns organisational interest. To deal with this problem, the similarity metric could be defined in a way that is sensitive to the query type. In this way, evidence obtained by sampling the sources may be better utilised in forming groups given the specific goal of a query.

The use of a learning interval in order to revise the similarity metric could lead to computation overheads, especially if there was no need for a revision. Model revision should rather be based on evidence obtained from interaction with the system. Where learning is not expected to lead to a significant revision, a decision-maker may instead adapt its sampling strategy in line with the current state of the model. In Chapter 5, we focus on learning of sampling strategies, and this sort of idea is taken into consideration.

4.5 Summary

In this chapter we presented a source diversification model, that allows a decision-maker to group sources based on their perceived similarity. The model uses evidence from the past reports of sources in order to learn a similarity metric. This metric is subsequently employed for stratifying the source population. We adopt techniques from machine learning in order to learn a generalisation from the notion of similarity in reports of sources to similarity in their features. In this way, unknown sources can be placed into groups, even when evidence of their behaviour in not available. The model is aimed at supporting a decision-maker to acquire as accurate an estimate as possible within budget limits.

The results of our experiments show the efficacy of our model in guiding reliable assessments. In particular, where hidden networks or patterns defining correlated behaviour exist in the population, our source diversification model is able to identify and exploit such structures in order to make better assessments of what is true in the world. While a naïve approach for truth discovery would perform poorly under these conditions, our model shows positive outcomes that outperform classical trust approaches in different experimental conditions.
Our model may be sensitive to different parameter settings. In particular, as demonstrated in our experiments (see Figure 4.8), DBS adapts in slightly different ways to different parameter values of the diversity threshold, $\psi$. Although this observation does not diminish the significance of our results, we believe that our model might benefit from appropriate parameter tuning to ensure optimal results.

We have identified the need to incorporate more robust decision-theoretic mechanism to handle complex source selection strategies. This will enable us to meet different information needs. For example, the cost and risk analysis of interacting with certain groups of sources may serve to inform how sampling decisions are made. The source diversification model presented in this chapter provides a good basis for driving such intelligent source selection strategies.

In the next chapter we demonstrate how effective sampling strategies can be learned by exploiting information from a diversity structure.
Chapter 5

Sampling Decision-Making

Until now, our focus has been on mechanisms by which a decision-maker can group sources based on their perceived similarity. The diversity model (see Section 4.2) is, however, only one part of an agent’s multi-source integration mechanism. The second necessary part, and the focus of this chapter, involves exploiting this model in sampling. That is, given a diversity structure, how can we optimally sample diverse groups of sources?

In Chapter 4, we adopted heuristics, commonly used in surveys, and trust for sampling. This entails sampling according to the proportion of sources in groups or the level of trustworthiness of the groups. By sampling proportionally according to the size of groups, an agent faces the risk of allocating unnecessary budget to large but unreliable groups with the possibility of biasing the decision-making process. On the other hand, allocating sampling resources solely on the basis of perceived trustworthiness of group members while necessary, is not a sufficient criteria in sampling decision-making. For example, additional samples drawn from groups with trustworthy sources may be superfluous, and not necessarily lead to a further improvement in the estimate. This may, for instance, be as a result of the high level of similarity among sources in the sampled groups. The goal of the decision-maker is to improve its estimate of ground truth with a limited set of samples. In other words, the agents must be able to optimally sample from diverse groups while maintaining the quality of information.

In this chapter we adopt a decision-theoretic approach to identify optimal sampling strategies. A decision-theoretic framework plays two significant roles. First, and most important, it enables a precise and concrete formulation of an agent’s source selection problem. Second, it provides mechanisms for designing a solution to the problem. An agent, with the aid of decision theory, can select effective sampling strategies based on its current trade-offs. For instance, the cost and risk analysis of interacting with certain groups of sources may serve to inform how sampling decisions are made. Utilising a decision-theoretic approach to guide sampling decisions will, however, require a richer representation of an agent’s sampling task than we have adopted so far. We extend our previous view of an agent’s task to include concepts such as sampling states and actions, that are relevant to our decision-theoretic model.

5.1 The DRIL Model

The Diversity modelling and ReInforcement Learning (DRIL) model for sampling decision-making is based on the idea of combining diversity modelling with reinforcement learning. Figure 5.1 provides an overview of the decision-making process, showing the relevant concepts that constitute
5.1. The DRIL Model

Figure 5.1: The DRIL model

an agent’s decision-theoretic approach to source selection.

In sampling the source population, a decision-maker has to make decisions regarding the number and the identity of sources to query in order to maximise its utility. When the agent takes a sampling action by querying a subset of sources, it receives a reward. This reward, which represents the utility for adopting the selected action or strategy is incorporated into the decision model in order to inform or refine subsequent decisions. For example, the decision-maker may be more inclined to adopt strategies that have been observed in the past to yield better pay-offs or rewards. The sampling state represents the agent’s current knowledge about the dynamics in the source population, which may in turn inform the type of strategy to implement. In particular, the concept of sampling states is captured by the degrees of agreement and trustworthiness of sources in groups within a diversity structure. A task constraint represents the relative preference of the decision-maker or the relative importance of the agent’s task. For instance, a highly-sensitive task may require the use of more resources in order to obtain the required level of confidence in the estimate. The decision model must be able to adapt to this and other preferences, as they might impact the reward received by the agent. Outcome of the estimation task is used to update both the agreement and trust models of the various groups in the diversity structure, and hence the sampling state.

We provide detailed description of the concepts making up the decision-making process. Later in this chapter, we will present a concrete realisation of the decision model, demonstrating how it can be utilised by an agent to guide sampling decisions.

5.1.1 Sampling State

The main problem in performing efficient sampling is that there is an inherent uncertainty about the behaviour of information sources in the environment. In particular, the decision-maker is uncertain about the correlation in the reports of sources or how reflective of ground truth they may turn out to be. If the decision-maker knew these dynamics, it could sample from diverse groups in a more
5.1. The DRIL Model

clever manner. For instance, the decision-maker could sample less from groups with very similar sources, or avoid the allocation of more resources to groups that do not guarantee a significant improvement in the estimate. One possibility is to estimate these dynamics while sampling from the groups, using the acquired knowledge to refine the sampling strategy (Carpentier and Munos, 2011).

A sampling state reflects knowledge available to a decision-maker about the dynamics of the environment, which, in the context of our framework, is the behaviour of sources in diverse groups.

**Definition 30 (Sampling State)** Let \( G_k \in DS \) denote a group in a diversity structure. Also, let \( \tau_k \) and \( \sigma_k \) denote the trust and similarity parameters of \( G_k \) respectively. A sampling state is a tuple \((T, \Sigma)\), where \( T = <\tau_1, \ldots, \tau_K> \) and \( \Sigma = <\sigma_1, \ldots, \sigma_K> \) are vectors corresponding to the trust and agreement levels of groups \( G_k \in DS, \forall 1 \leq k \leq K, K = |DS| \).

The decision-maker can update its estimate of the state parameters, hence the sampling state based on evidence (feedback) derived from the interactions with sources in diverse groups. Therefore, the system can be in one of several states at any point in time. We denote the set of possible states as \( S \), and sampling decision-making is based, at least in part, on this information.

5.1.2 Sampling Strategy

At each sampling time, a decision-maker must decide how to sample from the source population. A sampling strategy or action is an allocation model used in assigning sampling resources to groups in a diversity structure.

**Definition 31 (Sampling Strategy)** Let \( G_k \in DS \) denote a group in a diversity structure, and \( G_k = \{0, 1, \ldots, |G_k|\} \) be a finite set of possible sampling allocations to \( G_k \), \( |G_k| > 0 \). A sampling strategy is a tuple \((g_1, \ldots, g_K)\) s.t. \( g_k \in G_k \forall 1 \leq k \leq K, K = |DS| \).

This definition implies that an agent may decide to sample zero up to all the sources in a group. The decision-maker has at its disposal a finite set of sampling strategies or actions, \( \mathcal{A} \). The action space \( \mathcal{A} \), dependent on \( DS \), is given by the \( K \)-ary product over \( K \) sets \( G_1, \ldots, G_K \), which is the set of \( K \)-tuples:

\[
\mathcal{A} = \prod_{k=1}^{K} G_k
\]  

(5.1)

The action space represents all possible sampling combinations available to the decision-maker. The expression, \( \mathcal{A} = \{(0, 0, 0), (0, 0, 1), \ldots, (3, 2, 5)\} \), is an example action space based on a diversity structure with three groups, \( G_1, G_2, G_3 \in DS \), having sizes \( |G_1| = 3, |G_2| = 2, |G_3| = 5 \), and \( |\mathcal{A}| = 72 \). An action, \((2, 0, 1)\), implies that 2 sources are sampled from the first group, \( G_1 \), 0 from the second group, \( G_2 \), and 1 from the third group, \( G_3 \). The tuple \((0, 0, 0)\) is the null action, where no sources are queried.

In settings where there is a limitation, \( \Phi \), in the number of sources to sample, the sampling strategy is defined such that:

\[
\sum_{k=1}^{K} g_k \leq \Phi.
\]  

(5.2)
The sampling action adopted by an agent implicitly encapsulates the notions of cost and risk of a transaction. For example, an agent that decides to sample more information sources from the population, might be actively making an investment commitment. This commitment might be necessitated by the high level of risk associated with the agent’s sensing task. It might well be that by adopting a strategy where more information sources are selected, or approaching specific groups of sources, the agent is expecting to achieve a high-quality estimate. Executing an action has two effects. First, it yields some pay-off or reward for the decision-maker. Secondly, it provides information (or feedback) that may be used to update the agent’s knowledge about its environment. The desirability of an action depends on a number of factors including the sampling state, expected pay-off, and so on.

5.1.3 Reward

In making sampling decisions, the decision-maker aims at obtaining as good an outcome as possible. A reward is a feedback from the system, that reflects the “goodness” or the relative desirability of the outcome of an agent’s action. The effect of an agent’s sampling decision depends not only on its choice of action, but also on factors that may be outside of the decision-maker’s control (e.g., the behaviour of sources in the system or more generally the sampling states). For instance, if the decision-maker assigns a large sampling budget to a certain group, and reports sampled from the group turned out to be very similar, then the agent would have wasted resources through redundant sampling. The effect of this strategy is ultimately reflected in the reward received by the agent. The possible outcomes, and hence the rewards received by an agent for executing an action are defined as the combined effect of the chosen action and the sampling state.

**Definition 32 (Reward)** Let $S$ denote the set of sampling states, and $A$ the set of actions. We define a reward as a function $r: S \times A \rightarrow \mathbb{R}$.

Generally speaking, numerically represented rewards (or utilities) are easy to use in decision-making. An intuitive decision-rule is to choose the alternative with the highest reward. However, if there are more than two alternatives with maximal value, one of them must be selected. This is often referred to as the *rule of maximisation* (Hansson, 2005). We assume rewards to be both independent and additive.

5.1.4 Task Constraint

In Section 4.1.3, we defined sampling utility, $u$, as a function of information quality and sampling cost. These two objectives intuitively capture the constraints under which a decision-maker’s task may be carried out. In particular, the decision-maker has a dual objective of maximising both the quality of information and the negative cost of sampling.

In single-criterion decision-making, the goal is to find a solution which is the best for the problem. For instance, the goal might be to find the best subset of sources to query in order to achieve the highest information quality possible, or finding the most cost-effective subset with which to interact. In other words, the objective may be to either maximise quality or minimise cost, but not both. In a multi-criteria decision-making, the key is to find compromising solutions that balance both objectives (i.e., information quality and sampling cost). In this sort of decision problem, there is usually no unique, perfect solution. Rather, the goal is to find a set of trade-off optimal solutions referred to as the *Pareto optimal solutions* (Marler and Arora, 2004). In certain
5.2 A Realisation: TIDY₁

In Chapter 4, we presented the TIDY framework, and went on to describe a specific realisation of the framework (TIDY₀) that deals with source diversification. TIDY₁, our realisation of TIDY in this chapter is concerned with learning sampling strategies. In doing so, we build on the techniques described in Chapter 3. In particular, our decision model is based on Reinforcement learning, RL (see Section 3.4).

5.2.1 Learning Sampling Strategies

Given a set of sampling actions, we expect a decision-maker to select a strategy that is most suitable for its current situation or preference. Selecting between different strategies would be trivial if the decision-maker could observe and characterise its environment (or the behaviour of sources) perfectly. However, this is rarely the case. Most often the agent’s environment is shrouded with uncertainty. Therefore, we need to provide the decision-maker with the means to learn appropriate source selection or sampling strategies in order to increase overall value of information (Bisdikian et al., 2014).

A decision-maker can exploit RL, which has the merit of handling complex dynamic and delayed consequences of decisions, to identify good sampling strategies in a stochastic environment to maximise its reward. We provide a more detailed treatment of some of the concepts described above, and demonstrate how DRIL can be used for the purpose of learning sampling strategies. Such strategies must be able to make intelligent trade-offs between quality and cost of information. Our approach to this entails designing a multi-criteria reward function that takes both the quality of information and the cost of sampling into account to give rewards in reinforcement learning.

Learning with Multiple Criteria

The use of rewards to formalise an agent’s goal is one of the most distinctive features of RL (Sutton and Barto, 1998). Rewards received by an agent for implementing different sampling strategies can be used as a basis for learning efficient sampling strategies. This is the so called evaluative feedback in RL, that indicates how good an action is but not whether it is the best or the worst possible action.

Earlier, we defined a reward, \( r \), as a function that maps a set of sampling states, \( S \) and actions, \( A \) to a numeric signal or value. This definition may, in some ways, suggest that the agent’s aim is
to optimise a single objective that is expressed as a function of a scalar reward or reinforcement. This is not necessarily the case. When dealing with a multi-criteria decision such as optimising both the quality and cost of sampling, instead of receiving a single scalar reward, an agent gets a reward vector representing both objectives to be achieved. That is, the single reward signal, \( r(s, a) \) \((s \in S, a \in A)\) is decomposed into a vector (Equation 5.3), where \( r_{\text{qual}} \) and \( r_{\text{cost}} \) represent rewards for the quality and the cost objectives respectively. This extension to traditional RL techniques is often referred to as multi-objective or multi-criteria RL (MORL) (Gábor et al., 1998; Vamplew et al., 2011).

\[ \bar{r}(s,a) = [r_{\text{qual}}(s,a), r_{\text{cost}}(s,a)] \] (5.3)

Learning optimal sampling strategies depends on the ability to learn in the presence of multiple rewards. A popular approach to solving MORL is to transform a multi-objective problem into a single-objective problem by employing scalarisation functions. The primary advantages of this approach is the reduced computational cost and reduced time spent in interacting with the environment, thus making scalarisation approaches more scalable in real online learning systems (Vamplew et al., 2011). Most scalarisation functions define an aggregate one-dimensional utility function by taking a weighted sum of the various rewards (Moffaert et al., 2013; Legriel et al., 2010). We adopt a similar view in this research, and define the reward function:

\[ r_{\text{qual, cost}} = \lambda r_{\text{qual}} + (1 - \lambda) r_{\text{cost}} \] (5.4)

The parameter, \( \lambda \in [0,1] \), is a coefficient that allows us to control the trade-off between the rewards, and intuitively captures the notion of a task constraint in our decision model. A decision-maker may, for instance, place higher (or lower) weights on an objective to emphasise its relative importance in a monitoring task. This of course means that we can redefine our sampling utility, \( u \) from Section 4.1.3 to incorporate the task constraint, \( \lambda \):

\[ u : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R} \] (5.5)

**Example 3**

To illustrate this concept, consider four sampling strategies: \( a_1, a_2, a_3, a_4 \in A \), and a reward vector of the form \( \tilde{r} = [r_{\text{qual}}, r_{\text{cost}}] \). The rewards for the strategies are: \( r_{a_1} = [0.6, 0.6] \), \( r_{a_2} = [0.7, 0.9] \), \( r_{a_3} = [0.8, 0.1] \), \( r_{a_4} = [0.7, 0.1] \). Further, assume the following coefficients: \( \lambda = 0.9 \) and \( \lambda = 0.1 \), representing the relative preferences of a decision-maker at different instances.

Table 5.1 shows the derived utilities for the different actions based on the different preferences. The constraint, \( \lambda = 0.9 \), suggests that the decision-maker has a high preference for quality, whereas, \( \lambda = 0.1 \), suggests a high preference for cost-minimisation. The utilities in each case are obtained using Equation 5.4. For example, the decision-maker derives a utility of 0.73 for implementing sampling strategy \( a_3 \) under a high quality constraint (i.e., \( \lambda = 0.9 \)). Implementing the said strategy, involves the use of more resources as suggested by the low reward for the cost objective \( r_{\text{cost}} = 0.1 \). It turns out, however, that by achieving a high quality level with strategy \( a_3 \), the
5.2. A Realisation: TIDY

<table>
<thead>
<tr>
<th>$r_{a_1}^*$</th>
<th>$r_{a_2}^*$</th>
<th>$r_{a_3}^*$</th>
<th>$r_{a_4}^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[0.6, 0.6]$</td>
<td>$[0.7, 0.9]$</td>
<td>$[0.8, 0.1]$</td>
<td>$[0.7, 0.1]$</td>
</tr>
<tr>
<td>$\lambda = 0.9$</td>
<td>0.6</td>
<td>0.72</td>
<td>0.73</td>
</tr>
<tr>
<td>$\lambda = 0.1$</td>
<td>0.6</td>
<td>0.88</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Table 5.1: Utility table for sampling example

![Figure 5.2: Reward relation in sampling example](image)

A decision-maker derives a higher utility as compared to, say, strategy $a_2$ with a higher reward for the cost objective, but a relatively lower quality level. In addition, we observe the effect of different preferences on the desirability of an action. For instance, when there is a constraint on quality (e.g., in highly-sensitive sensing tasks), the strategy $a_3$ yields more utility than $a_2$. However, when the preference is on cost-saving rather than on quality (e.g., in a low-risk task), strategy $a_2$ clearly becomes a more favourable choice of action to adopt.

Although there are four possible sampling strategies (with rewards shown in Figure 5.2), only two of them are optimal for any values of $\lambda$. One strategy occurs when $\lambda < 0.89$. In this case, the decision-maker should select strategy $a_2$, because its utility otherwise would be no higher than $u(0.7, 0.9, \lambda)$. The second case occurs when $\lambda \geq 0.89$. In this instance, the decision-maker should select strategy $a_3$, since its utility in taking any other action will be no higher than $u(0.8, 0.1, \lambda)$. An optimal agent would avoid selecting strategies $a_1$ and $a_4$ (represented as points 1 and 4 in Figure 5.2). No matter the task constraints or the decision-maker’s preferences, there exists some other strategy (i.e., points 2 or 3 in Figure 5.2) that dominates. Generally speaking, points 2 and 3 in Figure 5.2 are not strictly dominated by any other, and thus are considered to be Pareto optimal. Our aim is to use RL techniques to guide sampling decisions by selecting those strategies that are viable under different task constraints.
5.2. A Realisation: TIDY

Learning in Stochastic Environments

Generally, it is assumed that the RL agent has no prior knowledge of the environment except knowing the valid choice of actions. By trying alternative sampling actions, each characterised by a probability distribution over rewards that is initially unknown, the agent gains information about the environment which can be used to improve its strategy. A key factor in this context is the ability to maintain a balance between knowledge gained, and trying alternative action choices (March, 1991). The problem is further complicated in non-deterministic environments, where the actual utility of an agent’s choice is seldom known. Instead, what we have are estimates of an action’s value, which may not represent the true underlying distribution. For example, we may have imperfect sensors or sources which sometimes fail to provide accurate reports or whose reports may sometimes not reflect their true behaviour. The actual worth of interacting with these sources may not be revealed with just a single or a few interactions. The same parallel can be drawn in the choice of a sampling strategy, in that taking the same sampling action may lead to different utilities. In order to make good sampling decisions, the decision-maker must have access to reasonably accurate estimates of the expected utilities of its actions.

In the absence of a perfect knowledge of the environment (captured in RL context in terms of the reward and next state probability distributions), an agent must learn the long-run values of sampling actions without relying explicitly on a model. We employ temporal-difference (TD) or model-free learning, which is a well-known RL method to estimate the value of different sampling strategies. In contrast to model-based learning (see Section 3.4.3), temporal-difference learning assumes no model of the environment. The decision-maker uses experience from the execution of different sampling actions to update estimates of utilities, and hence to select desirable sampling strategies. For instance, the decision-maker may have learned from experience that the best strategy is to avoid sampling from certain groups of sources in certain task contexts. Such groups may, for example, comprise of low-grade sensors, the interaction with which may adversely impact the agent’s utility. The representative TD learning algorithm we adopt is SARSA (see Section 3.4.3), a simple and efficient model-free RL algorithm for learning a policy. SARSA maintains an estimate of the utility derived from implementing a strategy or action, $a$ when in state, $s$ encoded by a $Q$-value function $Q(s,a)$. In the case of a multi-criteria decision problem, scalarised $Q$-values denoted as $SQ$-values can be obtained by applying the scalarisation function:

$$SQ(s,a) = \lambda Q(s,a,r_{\text{qual}}) + (1-\lambda)Q(s,a,r_{\text{cost}})$$  \hspace{1cm} (5.6)

A $Q$-value is maintained for each of the objectives, $r_{\text{qual}}$ and $r_{\text{cost}}$, and action selection is carried out based on these $SQ$-values. We employ the Boltzmann action selection strategy as it is a more principled approach for exploration than other action selection techniques described in Section 3.4.4.

5.2.2 State Space Approximation

We now describe how an agent can compute $T$ and $\Sigma$ in order to represent its state space. We assume discrete (in fact binary) values for the group parameters, $\tau$ and $\sigma$, in order to use standard RL algorithms. Therefore, $\tau \in \{1, 0\}$, and $\sigma \in \{1, 0\}$, where a value of 1 for $\tau$ and $\sigma$ denotes trustworthy and agreement respectively. Similarly, a value of 0 denotes untrustworthy and disagreement.
respectively.

Since we assume discrete states, we define one possible function that maps continuous values of trust and agreement of groups to discrete state values. Before proceeding, we describe how evidence about the trustworthiness and agreement of groups are obtained.

At the end of each time step $t$ (i.e., after fusion and decision-making utilising $\hat{\theta}'$), we assume that the decision-maker can observe ground truth, or fact, $\theta'$. Positive, $r_{k:τ}$ and negative, $s_{k:τ}$ evidence about the trustworthiness of a group, $k$ can be obtained using the trust evaluation function given a reliability threshold value:

$$v_{\text{trust}}(\hat{\theta}'_k, \theta') = \begin{cases} (1, 0), & \text{if } |\hat{\theta}'_k - \theta'| \leq 0.1 \\ (0, 1), & \text{otherwise} \end{cases} \quad (5.7)$$

The method of computing the group estimate, $\hat{\theta}'_k$, is given in Equation 4.7. If a group has not been sampled in the time step $t$ (i.e., if $|g_k| = 0$ in the selected action), then its $\langle r_{k:τ}, s_{k:τ} \rangle$ pair at time $t + 1$ remains unchanged. This observation has some significant implications especially in dynamic systems. For instance, it reinforces the need for an effective exploration of the source population both in order to accurately estimate the behaviour of sources and to properly allocate sampling resources (Carpentier and Munos, 2011). The $\langle r_{k:τ}, s_{k:τ} \rangle$ pair is used to derive an expected trust value, $ω_{k:τ}$ using Equation 3.10 (see Section 3.1.4).

We can also compute the expected agreement value, $ω_{k:σ}$ of sources in a group, $k$. Reports obtained from sources sampled in a group are used as evidence for computing their agreement. Therefore, $ω_{k:σ} = 1$ for any group with $|G_k| = 1$, since there is absolute certainty about a source agreeing with itself. For groups with $|G_k| > 1$, and $|g_k| = 1$ in the chosen action, $\langle r_{k:σ}, s_{k:σ} \rangle = (0, 0)$. This is because evidence from more than one source is required in order to effectively estimate agreement. In all other cases ($\forall x, y \in |g_k|$) we update $r_{k:σ}$ and $s_{k:σ}$ by applying the agreement evaluation function as in Equation 4.1.

Having illustrated how evidence about the trustworthiness and agreement of groups are obtained, we are now equipped with relevant information for deriving parameters of the state space. In particular, associated with each group $G_k \in DS$ are the expected trust, $ω_{k:τ}$ and agreement, $ω_{k:σ}$ values computed using the sequence of observations until time step $t$. We then carry out the following mappings:

$$ω_{k:σ} = \begin{cases} 1 & : ω_{k:σ} > 0.5 \\ 0 & : ω_{k:σ} \leq 0.5. \end{cases} \quad (5.8)$$

$$ω_{k:τ} = \begin{cases} 1 & : ω_{k:τ} > 0.5 \\ 0 & : ω_{k:τ} \leq 0.5. \end{cases} \quad (5.9)$$
5.3. Evaluation

5.2.3 Reward Computation

The reward signal is computed at the end of every time step using the outcome of the agent’s action. According to Equation 5.4, the reward function is of the form: 

\[ r_{\text{qual}, \text{cost}} = \lambda r_{\text{qual}} + (1 - \lambda) r_{\text{cost}}. \]

The quality component, \( r_{\text{qual}} \) is based on the system’s estimate of the environmental state:

\[ r_{\text{qual}} = 1 - \left| \frac{\hat{\Theta} - \Theta}{\Theta} \right|, \quad \Theta \neq 0 \]  

(5.10)

The method of computing \( \hat{\Theta} \), is given in Equation 4.8. The cost component, \( r_{\text{cost}} \), is based on the total cost of all sampled source:

\[ r_{\text{cost}} = \frac{\text{cost}(N') - \text{cost}(N)}{\text{cost}(N')} \]  

(5.11)

In our experiments, we assume a unit cost for all the sources. Although our model can accommodate variable costs, we are of the view that this cost value is more informative, and allows us to demonstrate how performance varies with task constraints.

5.2.4 Fusion Set Formation

The fusion set comprises of a subset of sources, \( N \subseteq \mathcal{N} \), and the reports from these sources may be used to derive an estimate of the environmental state. A sampling strategy only specifies the “formula” for sampling diverse groups in a diversity structure. However, it does not specify exactly how sources are selected from the designated groups given the allocation. For example, the sampling strategy or action \((3, 0, 5)\) represents allocations to three groups, \( G_1, G_2, G_3 \in DS \), with 3 sources sampled from \( G_1 \), 0 from \( G_2 \), and 5 from \( G_3 \). The fusion set, \(|N| = 8\) (i.e., \(|N| = 3 + 0 + 5\)), in this case is formed by randomly selecting from each \( G_k \in DS \). In particular, each member of a group has the same probability of being selected, and we randomly select (without replacement) the number of sources queried for information as specified by the sampling strategy.

5.3 Evaluation

We empirically evaluate our decision-making model through simulations to demonstrate its effectiveness in guiding sampling decisions. We show that under different task constraints, sampling decisions that exploit a combination of trust and diversity can be beneficial and can significantly increase overall value of information to a decision-maker. In evaluating our approach, we focus on the following metrics: (i) the total reward or utility accumulated by the decision-maker up to a certain time-point; (ii) the corresponding accuracy of estimates with respect to ground truth.

The following hypotheses form the basis of our evaluation:

- **Hypothesis 1**: With high constraints on the cost of information acquisition under varying degrees of malicious sources, agents that actively learn sampling strategies perform better (in terms of utility or reward) than those that do not.

- **Hypothesis 2**: Under varying degrees of malicious sources, agents that combine source diversification and reinforcement learning to drive sampling strategy perform better at managing the trade-off between quality and cost of information than those that do not.

To test these hypotheses, we compare DRIL, which uses the realisation of the TIDY framework, TIDY$_1$, defined above to the following models of truth discovery:
### Table 5.2: Experimental parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta$</td>
<td>1.0</td>
<td>Learning rate (RL)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.1</td>
<td>Discount factor (RL)</td>
</tr>
<tr>
<td>$T_{mp}$</td>
<td>0.1</td>
<td>Temperature (RL)</td>
</tr>
</tbody>
</table>

**ADaptive Stratified Sampling (ADSS)**  Adaptive stratified sampling for optimal allocation is an instance of active learning (Cohn et al., 1996), with a number of practical applications in areas such as quality control, clinical trials, surveys, etc. In ADSS, the decision-maker estimates the ground truth by estimating the mean reports of each stratum (or group), and then combines the estimates to obtain an overall estimate. The general focus here is to sample those groups (in a diversity structure) that have high report variance. Since the allocation is based on estimates of variance, the ADSS agent at each time step learns estimates of the report variance in the different groups. These estimates are then used as a basis for allocating sampling resources to groups. Thus, if sampling more from a group does not reveal additional information (or leads to very similar reports), an agent, in the next time step would decrease the allocation to the group in order to save cost. While ADSS approaches focus on allocating sampling resources in order to reduce the variance in the final estimate, they generally do not take the trustworthiness of sources into account in making sampling decisions. Notable research in this area include Etoré and Jourdain (2010); Carpentier and Munos (2011); Kawai (2010); Antos et al. (2010). By comparing our model with ADSS, we can assess the value of a strategy that is not only sensitive to the variation of reports in groups, but also uses the relative trustworthiness of the sources to guide source selection.

**Classical Trust and Reputation Sampling (CTRS)**  Classical trust and reputation approaches focus on accurately estimating the quality of information from sources rather than optimally selecting the sources to query. This strategy, adopted by trust and reputation models, usually samples all sources in the population, and either follows majority opinion, or uses some discounting method given estimated trustworthiness of the sources. Notable research in this area include Jøsang and Ismail (2002), Regan et al. (2006), and Teacy et al. (2006). Again, we model the trustworthiness of each source using a Beta probability density function. In estimating the environmental state, trust assessments are used to discount reports received during fusion. Comparing our model with CTRS allows us to assess the value of strategic sampling of information sources to the overall value of information to a decision-maker especially in environments where there is a concomitant cost to sampling. The complexity and high computation costs associated with other proposed models such as HABIT (Teacy et al., 2012) make them unsuitable as benchmarks in the context of this research.

### 5.3.1 Experimental Environment

In evaluating our approach, we utilise the multi-agent simulation environment presented in Section 4.3.1. In this section, therefore, we only describe aspects of the experimental environment and parameters that are different from those in Section 4.3.1.
Reliable reports are drawn from a Gaussian distribution, \( N(\theta, 0.01) \), while malicious reports follow \( N(\theta + \varepsilon, 0.01), \varepsilon \in [0, 5] \). In order to focus on learning sampling strategies, we assume a fixed diversity structure throughout the duration of an experiment. Thus, the diversity model is updated only once (after the initial learning interval, \( L \)). Other parameters, including those specific to reinforcement learning are captured in Table 5.2. In particular, the learning rate, \( \eta \), set between 0 and 1, determines to what extent newly acquired experience overrides older ones. A value of 0 means that the scalarised \( Q \) values, \( SQ(s,a) \), are never updated, hence nothing is learned. On the contrary, a high learning rate implies that learning can occur quickly. We set the learning rate, \( \eta \) at 1.0 to give full weight to new experiences. The discount factor, \( \gamma \) is also set between 0 and 1. This parameter controls to what extent expected future rewards affect current decisions (see Section 3.4.2). As the agent isn’t expected to be aware of the duration of tasks beforehand, we set value of this parameter at 0.1, in order to place more weight on immediate rewards than expected future rewards. Finally, the temperature, \( T_{mp} \) parameter controls the amount of exploration in the learning process (see Section 3.4.4). Setting the value of this parameter at 0.1 allows us to bias action selection towards those with high \( SQ(s,a) \) estimates.

### 5.3.2 Results

We present and analyse the total reward (utility) and the mean absolute error (accuracy) averaged over multiple runs for the different strategies considered, under different task constraints. Each instance of our simulation was repeated 20 times. Although we considered an infinite-horizon monitoring task, the system was observed to have stabilised by 100,000 time steps. Statistical significance of differences between strategies was computed using analysis of variance (ANOVA). Analyses of significant differences between pairs of strategies were carried out using Tukey’s HSD (\( \alpha \)-level = 0.05). We present and analyse the reward and mean absolute error (information quality) averaged over multiple runs for the different strategies considered. Error bars represent a 95% confidence interval (CI) variation between means.

**Hypothesis 1**

In our first set of experiments, we assess the value of actively learning sampling strategies to an agent whose interest is to minimise costs. We ran our simulation with task constraint values \( \lambda \) of 0.1 and 0.2, which represent high preference for cost-minimisation. Under these experimental conditions, the decision-maker is rewarded more for implementing strategies that lead to the use of fewer resources. In real-world contexts, higher preferences for cost-minimisation may reflect transactions or tasks associated with lower risk to the decision-maker. For instance, sampling a huge number of sensors at certain times of the year (e.g., during summer months, with mostly sunny spells), in order to learn about the state of the environment in a flood risk area may be counter-intuitive. This is because the risk of flooding at such times is very low. Therefore, the cost of sampling that many sources may outweigh the value derived. In some other setting, a buyer in an electronic marketplace may be unwilling to invest huge effort sampling for information about the reputation of a potential seller, especially if the transaction involves an insignificant amount of money. The costs (say, in time or bandwidth) in dealing with multiple requests, may outweigh any potential knowledge about the expected behaviour of the seller regarding that particular transaction.

The null and alternative hypotheses are:
5.3. Evaluation

- $H_0$: There is NO significant difference (two-tailed) in the performance of agents that actively learn sampling strategies and those that do not.

- $H_1$: There is a significant difference (two-tailed) in the performance of agents that actively learn sampling strategies and those that do not.

The result of our simulation is shown in Figures 5.3 and 5.4. Both DRIL and the adaptive sampling strategy (ADSS) employ learning mechanisms in order to decide how to sample the source population. These approaches as observed in Figures 5.3 (a) and 5.4 (a) outperform the classical trust and reputation strategy (CTRS) in terms of the utility or reward. Unlike DRIL and ADSS, CTRS does not actively learn sampling strategies. As the focus of CTRS is mainly on maximising quality, this approach samples all the sources in order to achieve this objective. Therefore, in
settings where emphasis is less on quality than cost-saving, CTRS-based approaches are bound to perform poorly in terms of utility.

As expected in Figures 5.3 (b) and 5.4 (b), CTRS performs better than DRIL and ADSS in terms of estimation accuracy (mean absolute error). Sampling all the sources gives CTRS an advantage in this context. This seeming advantage in quality, as observed in Figures 5.3 (a) and 5.4 (a), is insufficient to offset the impact on the agent’s reward, which is due to the non-optimal sampling decisions made by CTRS. DRIL on the other hand uses reinforcements from the system to adapt its sampling strategies towards actions that lead to better pay-offs. In particular, DRIL adopts sampling strategies that involve the selection of fewer sources from diverse groups. However, in so doing, the approach still maintains a relatively good balance in quality. The results in Figures 5.3 (b) and 5.4 (b) show that the performance of CTRS in terms of quality is similar to that of DRIL. The slight improvement observed in the performance of DRIL in Figure 5.4 (b) can be attributed to the change in the value of $\lambda$. In particular, increasing the value of $\lambda$ from 0.1 to 0.2 implies that less constraint is placed on cost. DRIL therefore exploits this situation to sample more sources, thus improving the accuracy of its estimates. Although the preference in this context is not on quality, this result is actually encouraging. That is, in the extreme case when emphasis is on cost-minimisation, the DRIL model can lead to much higher utility than CTRS-based approaches. At the same time, the approach is able to maintain the quality of information.

The results are found to be statistically significant, and support our hypothesis that learning sampling strategies can minimise cost of information acquisition, thereby yielding better pay-offs to a decision-maker. In particular, the analysis for Figure 5.3 (a) shows that $p = 9.28 \times 10^{-38}$. A post hoc analysis suggests the utility obtained by DRIL is 7 times higher than that of CTRS, and the utility obtained by ADSS is 6 times higher than that of CTRS. The adjusted $p$-value in each case is $\ll 0.001$. Similarly, the analysis for Figure 5.4 (a) shows that $p = 1.36 \times 10^{-32}$. A post hoc analysis suggests that the utility of DRIL and ADSS are again, higher than that of CTRS. The slight drop in performance in this later case, is due to the fact that slightly less emphasis is placed on cost when $\lambda = 0.2$. It is however, not immediately apparent if DRIL has a superior sampling strategy to ADSS. There is no significant difference in the rewards obtained by both approaches in both instances ($\lambda = 0.2$, $\lambda = 0.2$). The learning mechanisms employed by these approaches exploit the stratification of the source population to develop their respective sampling strategies. The results in Figures 5.3 (b) and 5.4 (b) are also found to be statistically significant. The analysis for Figure 5.3 (b) shows that $p = 0.017$. A post hoc analysis suggests that the only group difference in this instance is between CTRS and ADSS, with an adjusted $p$-value of 0.019. The analysis for Figure 5.4 (b) shows that $p = 0.014$. A post hoc analysis suggests both DRIL and CTRS outperform ADSS. However, there is no significant performance difference between DRIL and CTRS.

**Hypothesis 2**

Trust is often regarded as an effective component for truth discovery, especially in environments where there is uncertainty regarding the reliability of sources. Trust may be a necessary component in many interaction contexts (Moorman et al., 1992; Rousseau et al., 1998), however, it might not be a sufficient criteria for making sampling decisions. As discussed in Section 5.1.1, developing an optimal sampling strategy may involve identifying and exploiting other dynamics in the
environment. For instance, taking the diversity in the source population into consideration when
developing sampling strategies, has earlier been demonstrated to be beneficial to a decision-maker,
especially in contexts with a high value on cost-saving.

![Figure 5.5](image1)

**Figure 5.5:** Reward and mean absolute error for $\lambda = 0.3$

![Figure 5.6](image2)

**Figure 5.6:** Reward and mean absolute error for $\lambda = 0.4$

In this set of experiments, we evaluate the value of combining both diversity and trust within
a decision-theoretic context, and how this impacts the ability of a decision-maker to manage dif-
ferent trade-offs between quality and cost. We ran simulations with different task constraint values
($\lambda$). By this, we aim to represent the different levels of preferences that may be important to the
decision-maker. We compare the performance of our decision model, DRIL, in these contexts
to ADSS, which develops its strategies solely on the basis of a stratified population, and CTRS,
whose strategy is based solely on trust.

The null and alternative hypotheses are:
• $H_0$: There is NO significant difference (two-tailed) in the performance of agents that combine source diversification and reinforcement learning to drive sampling strategy and those that do not.

• $H_1$: There is a significance difference (two-tailed) in the performance of agents that combine source diversification and reinforcement learning to drive sampling strategy and those that do not.

Figures 5.5 and 5.6 show the conditions when the task constraint, $\lambda$ is set at 0.3 and 0.4 respectively. Under these conditions, the preference leans slightly towards maximising quality. As observed in the graphs, DRIL performs better than CTRS and ADSS in terms of utility. A similar performance level is observed when the task constraint is set at 0.4, signifying a slight shift in the preference towards quality. Under this condition, however, there is an interesting change in the relative performances of the approaches. The reward obtained by CTRS slightly improves while that of ADSS shows a slight decline. This shift in pattern goes in line with the (slight) shift in emphasis from cost to quality. In this instance, the CTRS strategy is rewarded more than before for its accuracy, leading to an improved utility. ADSS on the other hand specialises in cost-saving by learning to sample fewer sources. The bulk of the pay-off obtained by this approach comes from effectively optimising the cost objective ($r_{cost}$). However, as emphasis is slightly shifted from cost, the pay-off for this approach also experiences a (slight) negative impact.

Intuitively, one may assume that CTRS performs better than the other approaches with regards to accuracy (or quality). This is because CTRS not only samples all the sources, its focus is mainly on modelling sources’ behaviour in order to assess the quality of information they provide. In both cases when the task constraint is 0.3 and 0.4 (see Figures 5.5 (b) and 5.6 (b)), CTRS is observed to perform much better than ADSS. An important deduction from these results is that, although ADSS obtains better utility (when compared to CTRS) under these conditions, the approach does poorly in maintaining the quality of information. DRIL on the other hand shows a more robust balance in maintaining the quality of information. This result is even more remarkable, given that in an attempt to obtain high utility, DRIL samples much fewer sources than CTRS. We explain this observation by the fact that DRIL uses feedback in reward signals together with information encoded in system states (similarity and trust in groups) to select high-rewarding actions. The statistical analysis for Figures 5.5 (b) and 5.6 (b)) suggest that the results are statistically significant, $p = 0.012$ and 0.01 respectively. While a post hoc analysis suggests that DRIL performs significantly better than ADSS in terms of information quality, there is no significant difference in the performances of CTRS and DRIL.

Figures 5.7, 5.8, and 5.9 show a pattern in the performances of the different approaches when the task constraint is 0.5, 0.6, and 0.7. In particular, DRIL continues to dominate the other approaches in all cases in terms of reward. The performance of CTRS in relation to DRIL shows a steady improvement in terms of reward with increasing value of $\lambda$. On the other hand, ADSS shows a steady decline in performance.
In terms of accuracy (mean absolute error), DRIL continues to perform very well. The difference between the performances of DRIL and CTRS becomes smaller with an increase in the value of $\lambda$. This is because DRIL is able to sample more sources in order to make relatively better assessments, as the constraint on cost is relaxed. In all the cases considered here (i.e., $\lambda = 0.5, 0.6, 0.7$), the statistical analyses in (under error) suggest that there is no significant difference in the performance of CTRS when compared to DRIL. This reinforces our earlier observation that DRIL is quite capable of maintaining the quality of information even while its strategy leads to much higher utility. Interestingly, the performance of CTRS as well as that of ADSS remains unchanged under these settings. This can be explained by the fact that, changes in task constraints do not necessarily affect the performances of both CTRS and ADSS strategies when considering only quality. The CTRS strategy does not adapt to changes in the decision-maker’s preferences, but continues to sample all the sources in the population. In the case of ADSS, its adaptive sampling strategy is driven mainly by local reinforcements involving the variance in the reports received from sources.
in different strata. The strategy does not learn based on the global reinforcement or reward involving both cost and quality. Therefore, given that changes in task constraints do not affect the dynamics of the sources themselves, this strategy continues to operate in a similar manner. By so doing, it obtains similar results (quality-wise) in all cases.

The next condition we consider, is when there is an extreme preference for the quality of information. In this setting, it is assumed that the decision-maker is willing to sacrifice cost in order to acquire better quality information. For example in a high-risk task, the decision-maker might be less concerned with the number of sources that must be queried than achieving a sufficiently accurate estimate. This is reflected in the higher weights ($\lambda = 0.8$ and $\lambda = 0.9$) placed on the quality objective ($r_{qual}$). Figures 5.10 and 5.11 show our results under this condition. First, we consider the case when $\lambda = 0.8$. While our DRIL model continues to perform better in terms of reward, the performance gap between CTRS and ADSS is much reduced. The statistical analysis (under reward) suggests that the performance of DRIL is significantly better than CTRS and ADSS, but there is no significant difference in the performance of ADSS over CTRS. With greater emphasis on quality, the savings made by the ADSS approach become insufficient to boost its reward. Through a careful selection of sampling strategies, however, DRIL is able to perform better than the other two approaches.

The performance of DRIL is challenged by the CTRS approach when the task constraint is 0.9. With a strict demand for quality, DRIL tends to adjust its sampling strategy towards sampling more sources in order to improve its likelihood of acquiring more accurate estimates. This is reflected in the very similar result obtained by both DRIL and CTRS in Figure 5.11 (b). However, in implementing a more greedy strategy biased towards quality, the clear advantage previously experienced by DRIL in terms of reward is diminished. The performance of CTRS is significantly improved in this context, which would be expected, given the high preference place on quality. The performance of ADSS deteriorates further in terms of reward. As shown in Figure 5.11 (a), ADSS only outperforms CTRS when the proportion of malicious sources is very low ($< 0.3$), after which CTRS dominates. Having said this, there is no significant difference in the performances of

Figure 5.9: Reward and mean absolute error for $\lambda = 0.7$
5.3. Evaluation

Figure 5.10: Reward and mean absolute error for $\lambda = 0.8$

After a careful analysis of sampling activity, subject to different task constraints or preferences, the results presented here support our hypothesis that a source diversification and reinforcement learning sampling strategy represented in our DRIL model, allows a decision-maker to better manage trade-offs between cost and quality. Specifically, the following conclusions are supported:

- In all task constraint cases, except for the extreme case where $\lambda \geq 0.9$, the DRIL model leads to a significantly better utility than approaches based only on source diversification or trust.

- In extreme cases, with a strict demand on quality $\lambda \geq 0.9$, DRIL performs significantly better than an approach based only on source diversification, and no worse than an approach based only on trust.
5.4 Discussion

Reasoning about the trustworthiness of potential information sources is important for truth discovery in large, open and dynamic systems. Reliance on trust alone is, however, insufficient for making effective sampling decisions. In resource-constrained environments with a concomitant cost to sampling, a decision-maker needs to devise means for optimally sampling the source population for evidence. Sampling all possible sources, an approach often adopted by trust and reputation mechanisms, may adversely affect the utility of the decision-maker: the cost of sampling some sources may outweigh the value derived. The decision model we have proposed allows a decision-maker to learn effective sampling strategies. Results of our evaluation show that, by utilising a combination of diversity modelling and trust within a decision-theoretic context, a decision-maker can better manage trade-offs between quality and cost of information than strategies based either only on trust or on stratification.

Our contribution in this chapter focuses on the application of reinforcement learning techniques to sampling decisions. In particular, we build our decision model around a diversity structure, and employ reinforcement learning to optimally allocate sampling resources to diverse groups. While RL provides a principled means to learn optimal sampling strategies, there are known complexity issues (Kakade, 2003). Particularly, the use of RL when no groups are formed (i.e., when sources are treated as individuals) is computationally expensive. For instance, in a similar environment as that described in our evaluation with 100 information sources, the learning algorithm would have $2^{100}$ different actions to select from in each state. Such a large action space would make it almost impracticable to choose a “good” strategy among alternative choices, thus defeating the purpose of our overall framework. Exploiting dependencies among sources in diversity modelling to form groups (see Chapter 4) enables this complexity to be managed. Also, as highlighted in Section 5.1.2, the action space can be further reduced in settings where budgetary bounds are imposed. The approximation of the state space in our evaluation further allows us to manage the size and complexity of the RL problem. As shown in our results, this approximation does not affect the model’s ability to make good source selection decisions, especially when compared to classical trust models without any behaviour approximations. In addition to the issues already advanced when considering the use of RL-only settings (i.e., without the formation of groups), we note that sources may collude or be biased in other ways. This increases the risk of sampling at greater cost, with possibly no improvement in estimates, thus making the use of a diversified approach in an RL context even the more pertinent.

As discussed in Chapter 2, active learning provides an important body of work useful for dealing with the problem of optimal sampling (Settles, 2009; Cohn et al., 1996). Under this area, the use of adaptive (optimal) allocation mechanisms has been advocated in the literature by a number of authors, including Etoré and Jourdain (2010). While there is much promise in this research in the context of our problem, they still fall short given that they generally don’t consider the problems of unreliability among information sources, and therefore don’t incorporate trust mechanisms. In the sort of environment we consider in this research where sources have the tendency to behave unreliably, these approaches fall short. That notwithstanding, the performance edge shown by a representative approach of adaptive sampling over trust-only approaches emphasises the need for learning strategies that would enable the intelligent sampling of the source population.
It is important to take into consideration the dynamic environment in which agents operate, and how that might impact our learning process. For instance, agents may join and leave the environment. Behaviours may change, implying that agents identified with certain groups, may, over the course of time need to be re-assigned. In general, updating the diversity model to reflect changes in the environment would be realistic if not a useful activity. Updating a diversity structure in the context of learning sampling strategies, may affect the learning process itself. For instance, each modification of the grouping structure may involve re-instantiating the RL process in terms of the state space and action space. Etoré et al. (2011) describe a strategy that enables sampling according to the variation in diverse groups, while at the same time adapting the group structure online. This is a non-trivial problem in itself. In order to focus on learning sampling strategies, we have not considered this effect in this research. We note, however, that while the group affiliation of an individual source may change, it may be the case that groups themselves are relatively stable in many domains (e.g., organisations tend to be more stable than individual memberships).

We believe that it would be useful to reason about the time it takes for the reinforcement learning algorithm to stabilise. The cut-off point of 100,000 timesteps used in our evaluation might be regarded high in many domains (e.g., military applications). That being the case, the system is able to stabilise much faster as information about group behaviour is gathered and reused over time. Also, as earlier mentioned, in many domains group activity may be more stable, thus significantly reducing the time it takes to stabilise.

5.5 Summary

In this chapter we presented a decision-making model that allows a decision-maker to optimally sample a population of sources for evidence. We proposed DRIL, a model that combines source diversification and reinforcement learning to identify effective sampling strategies.

The results of our experiments demonstrate that our decision-making model performs as well as classical trust approaches in estimating ground truth. By sampling significantly fewer sources, DRIL significantly increases the value of information acquired from sources for a decision-maker. In this way we relax the assumption that querying sources is cost-free.

We believe that addressing this problem is important in environments such as sensor networks, where working within resource constraints is critical, and in social networks where dependencies among sources increase risks of biases.
Chapter 6

Discussion and Future Work

In this chapter, we reflect on our research, discuss the limitations of the model we have presented, and outline potential avenues for future work.

6.1 Discussion

The problem addressed in this research is that of truth discovery in environments characterised by uncertainty regarding the trustworthiness of information sources and constraints in resources. In Chapter 2, we provided a survey of related work in multi-source integration with a particular focus on information fusion and source selection, both of which are key components in the truth discovery process. Within this context, we reviewed existing relevant work in trust and reputation; useful mechanisms for addressing uncertainties and biases. We highlighted some of the shortcomings of these approaches in different aspects of the truth discovery process. In particular, we outlined their weaknesses with respect to robustness to biases and effectiveness under resource constraints.

The model we have presented is centred around the idea of a diversity structure, where a similarity metric is learned and subsequently used to cluster sources on the basis of their features. In this sense, our approach is quite similar to the stereotyping approach proposed by Burnett et al. (2010). This approach exploits correlations between behaviour and observable features of agents to form groups, but differs from our work in a number of key areas. In the first instance, the problem they address is that of task delegation to an agent. We observe that this may not necessarily fit within the context of multi-source integration. This is especially the case in situations where much may be gained by way of acquiring perspectives from multiple agents. Also, while the focus of their stereotypical learning is to estimate the likelihood of a source exhibiting a certain degree of trustworthiness, our focus in diversity is to estimate the likelihood of sources providing similar reports irrespective of their content (reliable or otherwise). As argued by Etuk et al. (2013b), diversity is not necessarily measured by similar levels of trustworthiness. It is possible that sources in different groups do portray similar levels of trustworthiness, for instance, when considering sources from different but equally reputable organisations. Recognising such subtle relationships has far reaching implications and quite desirable in the sort of context we base our work. For example, while different groups of sources may be equally trustworthy, it is possible that there are significant cost disparities between groups. Given that our source grouping is not based on trustworthiness, it is able to adapt to situations where knowledge of ground truth is not available to the decision-maker (Kamar et al., 2012). The approach presented in Burnett et al. (2010) can only gather relevant experience with which it may build its stereotypical model if knowledge of ground
truth is available to the evaluator. A similar issue, as highlighted in our discussion in Chapter 2, is faced with the approach proposed by Dong et al. (2009a). This approach also has a strong reliance on ground truth in order to learn dependencies between sources. This assumption makes the applicability of these and similar approaches problematic in environments where ground truth is either not available, or observations are significantly delayed.

Our sampling decision-making approach allows a decision-maker to identify and adopt effective sampling strategies that are robust to different task constraints. In this regard, our approach does not simply conform to the “wisdom of crowds effect”. Instead, we consider source selection as an optimisation problem, and seek to provide an agent with a principled means of maintaining good balance between costs and quality of information. In this context, the risk associated with a task or the information needs of a decision-maker are crucial elements for deciding how to sample for evidence. We note that while the approximation of the state space (i.e., the trustworthiness and agreement of groups) in our reinforcement learning, RL formulation might appear to be sacrificing useful information, this approach actually allows us to manage the size and complexity of the optimisation problem. As we have demonstrated in our evaluation in Chapter 5, approximating the state space in the manner we have does not affect the ability of our model to make good source selection decisions. Approaches, such as Irissappane et al. (2014), Kamar et al. (2012), and Teacy et al. (2008), that operate in a similar context tend to be overwhelmed by the explosion of the problem space. As a result, these approaches can only be applied to problems of a very small scale.

Since our work is situated in the context of dynamic environments, it is useful to reflect on how the dynamism in the source population might impact the ability of our model to learn effective sampling strategies. For instance, a strategy is proposed in Etoré et al. (2011) whereby a decision-maker can sample and at the same time adapt the underlying group structure online. While it is important to maintain a valid model of diversity, constantly adapting groups in the manner suggested in Etoré et al. (2011) may add unnecessary overheads to our model. For instance, each group restructuring may lead to the reformulation of various elements of our optimisation problem, thus prolonging the time and complexity of convergence. We note that while the group affiliation of an individual source may change, it may be the case that groups themselves are relatively stable in many domains (e.g., organisations tend to be more stable than individual memberships). This is in contrast to the sort of environment considered in Burnett et al. (2010), where groups are necessarily ad hoc in nature. In our case, learned sampling strategies across groups may themselves be quite robust to many kinds or levels of population changes.

Although we have assumed that groups are relatively stable, if the underlying feature-behaviour correlations change significantly, then it would be to the benefit of the decision-maker to develop a new similarity metric. For example, companies may merge, thereby leading to changes in beliefs or ethos. The ability of a decision-maker to react to changes in the environment is mediated by two parameters: a learning interval, $L$, and a forgetting factor (Jøsang and Ismail, 2002). The forgetting factor is used to discount evidence in an agent’s memory over time. This is important, as incorporating too much old evidence can lead to anomalies in the model, especially in changing populations. The $L$ parameter provides an opportunity to improve the model by incorporating fresh evidence. The approach discussed in Chapter 4 uses this learning interval to determine
when the model of diversity should be rebuilt. As discussed, this approach has its limitations. The first is that manually setting this parameter is insensitive to the dynamics in the population, and as such may lead to unnecessary overheads. This problem is addressed by our sampling decision-making approach, which is able to adapt sampling strategies accordingly to reflect changes in the system without necessarily having to re-learn a new structure.

A limitation of our model, with respect to group formation, is the restrictive assumption that sources will always exhibit the same kind of dependencies. This may not necessarily be the case. Affinity between sources may shift depending on the nature of the question posed. It may be the case that a decision-maker is better off learning and employing different similarity metrics that align with specific query types. For example, the manner in which a population is stratified when a question concerns the national interest might be different from an instance when the question has more of an academic bearing. Consequently, we believe that our model could have been better enriched by a reputational mechanism such that decision-makers could share their experiences on possible ways of stratification. A similar idea is discussed by Burnett et al. (2010), where agents may share stereotype models. Agents may request stereotypical reputation about a target agent from other agents who may be better informed or may have built a stereotype of the target. In our problem it may not simply be a matter of requesting information on how a particular source may be classified. Knowledge sharing in the context of source diversification may be more demanding. For example, a polling organisation may request information from other agents on how best to stratify a population with respect to a particular query. If certain stratification variables (or combinations of variables) are known to be informative in the specific context, such information may prove useful, especially if the requester lacks experience. However, we note that sharing metrics for creating a diversity structure might propagate biases of other forms. For instance, there could be a deliberate attempt by malicious agents to mislead a requester into forming uninformative structures for selfish aims. Nevertheless, having some sort of guidance for group formation may be useful for an inexperienced agent or in the early stages of the system lifetime, rather than simply working off singleton groups in the bootstrapping stage. This is particularly useful in large systems as recognised by Carpentier and Munos (2011).

### 6.2 Future Work

In this section, we describe some potential applications where source diversification and strategic sampling strategies may be employed for making reliable assessments of ground truth. Following this, we discuss a number of opportunities for future research.

#### 6.2.1 Quality-Enhanced Crowdsourcing

Recently, there has been a huge amount of interest in human computation: utilising human abilities to perform computational tasks that are difficult for computers alone to process (Quinn and Bederson, 2011; Law and Von Ahn, 2009). In this context, crowdsourcing has emerged as an effective paradigm for human-powered problem solving (Ouyang et al., 2014; Tran-Thanh et al., 2014; Karger et al., 2014; Kamar et al., 2012; Franklin et al., 2011; Brabham, 2008; Kittur et al., 2008). As a distributed problem-solving and production model (Brabham, 2008), crowdsourcing is an umbrella term for a variety of approaches that tap into the potential of a large and open crowd of people to solve a variety of problems (Geiger et al., 2011). Howe (2006) defines it as:
the act of taking a task that is traditionally performed by designated agents (e.g., employees within an organisation) and outsourcing it to the crowd in the form of an open call.

According to Howe, a crucial prerequisite in crowdsourcing is the use of an open call format, and a large network of potential workers or contributors.

Many instances can be cited in which large crowds of people contribute to the solution of a problem, the realisation of which is often through the combination of knowledge, opinions, skills, etc. Notable ones include: Wikipedia, an online encyclopaedia in which any reader can be a contributor (Kittur et al., 2007); Yahoo! Answers, a people-driven question-and-answer web resource; InnoCentive, an open innovation web resource that enables “challenge problems” to be solved by anyone for either professional recognition or financial benefits; and Ushahidi, an emergency coordination resource. Crowdsourcing has been applied to solve tasks such as image classification or labelling, video annotation, optical character recognition, language translation, product categorisation, and form data entry.

A large class of crowdsourcing applications focus on solving consensus tasks, the goal of which is to identify a hidden state of the world by collecting multiple assessments from human workers (Kamar et al., 2012; Mao et al., 2013). In this context, a crowdsourcing application (e.g., Amazon Mechanical Turk) consists of two groups of users: requesters and workers. The application usually exhibits a list of available tasks presented by requesters. Each task may be associated with some reward (usually monetary), and there is a set time period for task completion. A worker may select a task to execute from the list of available tasks. At the expiration of a task, associated submissions are aggregated in some manner (e.g., voting), and the workers rewarded. Besides monetary rewards, a worker may gain credibility when his submission is accepted by the requester (Yuen et al., 2011). Figure 6.1 illustrates an example crowdsourcing process. The goal of the system is to predict the correct answer of a given task, based on reports submitted by workers.

Several characteristics of crowdsourcing make the process of truth discovery in this application space a challenge. First, because crowdsourced tasks may be tedious, and the pay low, errors are common even among the well-intentioned workers (Karger et al., 2014). For instance, a worker may want to complete as many tasks as possible in order to gain more rewards. With such motive, less effort (or time) is usually committed to the completion of tasks. In extreme cases, the system may be flooded with ill-intentioned workers whose sole aim is to exploit the process. They do so either by submitting arbitrary answers in order to collect their fee, or with the intention of misleading the system or the task requester. In the latter case, the workers often need to “gang-up” (collude) in order to create more impact in the system. This problem is exacerbated by the fact that most crowdsourcing instances take place within the context of a social network. Workers may very much depend on their “social ties” in order to complete a task. Social influence plays a significant role in individual decision-making, and may affect the reports submitted by workers (Lorenz

---

2https://answers.yahoo.com/
4http://http://www.ushahidi.com
5https://www.mturk.com/mturk
et al., 2011). A manifestation of this is through copying or the revision of personal opinions to conform to the ideas or opinions of others (Qi et al., 2013; Dong et al., 2010). Also, in Sybil attacks (Douceur, 2002), a single malicious worker can present multiple (fake) identities in order to control a substantial fraction of the system. This can have a substantial impact on the output or result submitted by the system to the requester. Another interesting problem is that workers are neither persistent nor, in some cases, identifiable. Each round of tasks may very well be completed by workers that are completely new to the system, and who may never again be encountered. For instance, poor workers may easily “whitewash” themselves because acquiring a new identity is easy and without cost.

In the circumstances outlined above, achieving quality, or more specifically accuracy, in crowdsourcing will require a robust integration process that is not only capable of dealing with biases, but also well positioned to meet different information needs of task requesters. Much work has been done in motivating people to provide quality answers (e.g., using reputation mechanisms: see our discussion in Section 2.3). For example, in Amazon Mechanical Turk a worker who frequently submits “bad” reports can be blocked from accessing future tasks (Quinn and Bederson, 2011). However, as mentioned, an expelled worker can easily gain access to the system by adopting a new identity. Furthermore, the process of determining what is bad differs from one system to another, the majority of which use some form of voting to reach a consensus. The true answer is often not known at the point of collecting worker opinions, making it difficult to identify responses provided by malicious workers. Also, given that the worker population is large, anonymous, and in some cases transient, it is generally very difficult to build trust relationships with individual workers (Karger et al., 2014). As previously mentioned, malicious workers may, for instance, exploit anonymity in the system to try to maximise their financial gains by producing generic answers rather than making an effort to produce reliable contributions.

By using a source diversification approach, logical group structures or dependencies among workers can be revealed. Contributions can then be aggregated at the group level, using a diversity-based fusion approach, rather than doing so directly from individual workers. This can prevent crowdsourcing applications from being inappropriately dominated by dependent sources, hence
mitigating the problem of correlated biases. Also, since we model trust at a group level, the requester or decision-maker can easily overcome the challenges posed by the ephemerality of workers with respect to trust formation. For instance, a priori trust for unknown agents (e.g., new workers) can be estimated using the discovered features based on our diversity model. In this sense, our feature-based approach to source diversification can also prove useful, especially for dealing with issues of anonymity in the worker population. Several authors, including Şensoy et al. (2014), Ross et al. (2010), and Downs et al. (2010) have shown how worker features may be obtained even in contexts of anonymity. Also, as observed in (Burnett, 2011, p. 116), shared views among workers can provide useful evidence for inferring correlations. This also aligns with our source diversification approach.

Another significant application of our model in crowdsourcing is for the preselection of contributors based on specific task criteria or information needs. For instance, a specific group of workers may be targeted based on their specific attributes. iStockphoto,6 for example, tends to favour workers (photographers) that have uploaded sample pictures adjudged to meet a certain standard. Research and Development (R&D) tasks in InnoCentive may be better targeted at certain classes of workers. Usually, potential solvers or workers in this system are asked for attributes such as degrees earned, research interests and so on during registration (Brabham, 2008). These features can be exploited by our model to preselect sources meeting certain task preferences. In general, context-specific restrictions (e.g., interests, age, nationality, profession) can be used as the basis for worker selection, rather than obtaining opinions (from workers) without any form of restrictions or strategy. This sort of preselection activity has been shown to be desirable for the enhancement of the quality of contributions in crowdsourcing (Geiger et al., 2011). Not only that, it eliminates the waste of resources (e.g., time, money) in processing undesirable or sub-standard contributions.

A slightly different application domain worth mentioning is supporting intelligence analysts (e.g., in military operations). In this context, our model can be used within CISpaces, a Collaborative Intelligence Spaces framework aimed at the support of intelligence analysis within a coalition (Toniolo et al., 2014). Developed as part of the International Technology Alliance (ITA),7 a research program funded jointly by the UK Ministry of Defence (MoD) and the US Department of Defence (DOD), CISpaces provides an analyst with a virtual space for individual and collaborative sense-making. In particular, CISpaces supports a hypothesis-driven crowdsourcing service, enabling analysts to request for information from different contributors. This enables well-informed decisions to be made for successful military operations. Given that the framework relies on crowdsourcing as a means of improving situational awareness, it is also subject to the same kinds of issues outlined earlier. The first step for delivering accurate intelligence products is, therefore, to distinguish credible information from unreliable and biased reports, in order to minimise the risk of deception. Our model proves useful in this context. This is especially the case where information sources from a specific coalition partner might collude in order to give a collective and possibly biased view. By revealing the dependencies among these contributors (using source diversification), a more efficient process of integrating evidence or intelligence can be achieved.

6http://www.istockphoto.com
7https://www.usukita.org
Also, by providing effective sampling strategies, our model can better support decision-making in complex operations in time-constrained environments. Not only would relevant selection strategies be employed, our approach guarantees the selection of fewer sources (reducing decision time) whose diversity can provide better coverage of the problem space. In addition, strategic sampling can enable a specific groups of sources to be targeted for crowdsourced sensing. For example, an analyst may approach doctors in an area to understand whether there has been an unusual increase in illness, which may indicate contamination of a water supply.

### 6.2.2 Monitoring Critical Infrastructure

The monitoring of critical infrastructure continues to receive much interest in many countries due to the significance of these assets to national economies. According to a US report, the growing threat of international terrorism has renewed the US Government’s interest in infrastructure issues (Moteff and Parfomak, 2004). Critical infrastructure can be regarded as:

> systems and assets so vital to a nation that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters. (Moteff and Parfomak, 2004)

Bush (2003) posits that a nation’s critical infrastructure provides the foundation for national security, governance, economic vitality, and way of life. According to the author, their continued reliability, robustness, and resilience, creates a sense of confidence and forms an important part of a nation’s identity and purpose. These infrastructures, as enumerated by Moteff et al. (2003), include: telecommunications; electrical power systems; gas and oil storage and transportation; banking and finance; transportation; water supply systems; and emergency services (including medical, police, fire, and rescue). For instance, water distribution networks, which bring water to our taps are complex systems, with tremendous impact on our daily living. An accidental or malicious contamination introduced into such networks can have severe impact (Krause and Guestrin, 2007). In any case, key assets can be individual targets, the compromisation of which could result in large-scale casualties and property destruction (Bush, 2003).

Oil and gas products constitute major components of any nation’s energy scheme. The pipelines carrying these products are critical infrastructure for the effective transportation and distribution of these energy resources. Given their significance in a nation’s energy plan, oil and gas pipelines can easily become targets of sabotage and terrorist attacks. Moreover, pipelines can unexpectedly fail for many reasons, such as corrosion, cracking, and process upsets (e.g., incomplete fusion, improper repair welds, incomplete penetration) (El-Darymli et al., 2009). For instance, according to Nigerian Federal Government figures, there were more than 7,000 spills between the years 1970 and 2000, and more than 1,000 spill cases have been filed against Shell alone. Shell, which works in partnership with the Nigerian government in the Niger Delta, reports that 98% of all its oil spills are caused by vandalism, theft or sabotage.\(^8\) While repairing and securing this infrastructure require large investment of money and time, failure to implement such an agenda often results in significant consequences, including loss of life, ecological risk, severe interruptions in service, damage to adjacent infrastructure and buildings, and multi-million dollar

6.2. Future Work

Figure 6.2: Pipeline monitoring using sensor networks

repair bills (Stoianov et al., 2007). Furthermore, oil spills tarnish corporate reputations and erode their credit ratings, resulting in financial loss due to lost oil, punitive fines and clean-up costs. For example, as at September, 2010, the Gulf of Mexico oil spill was estimated to cost BP a total of $40 billion.\(^9\) This incident, which led to the loss of life of 11 workers, is claimed to be the worst environmental disaster in US history. For these reasons, stakeholders in the industry have strong incentives to monitor the state of this critical infrastructure in order to ensure smooth operations.

Sensor networks are increasingly being employed for pipeline monitoring (Jawhar et al., 2007; Stoianov et al., 2007). These devices have the capability to detect, localise and quantify anomalies in transmission pipelines. In the case of water transmission pipelines, sensors can be used to monitor water quality in transmission as well as observing the water level in sewer collectors (Stoianov et al., 2007). One of the key advantages of using wireless sensor networks for pipeline monitoring is that not only can they be placed near the object of interest, they can be deployed in harsh and remote terrains not easily accessible to humans (Heidemann et al., 2006). In addition to conventional sensor networks, privately owned sensors such as cameras, GPS (Global Positioning System) devices, and cell-phones are also capable of providing useful information in the monitoring process, in a move known as community sensing (Krause et al., 2008). Figure 6.2 illustrates an instance of using sensor networks for pipeline monitoring.

Monitoring pipelines under the circumstances described above poses some interesting challenges. First, the use of sensors, as discussed in Section 2.1.2, implies that the system faces the challenge of resource constraints. This not only necessitates the need to minimise the rate at which these sources are queried, it calls for a clever sampling strategy capable of providing the much needed coverage under resource constraints. Secondly, the social and political aspects associated with pipelines add a different dimension to the challenge of pipeline monitoring. For

instance, stakeholders may maintain different levels of trust amongst one another that dictate their information sharing policies. In such contexts of high mistrust involving stakeholders of various degrees of trust and intentions, information providers often resort to transformation of information (e.g., obfuscation) to manage the inferences that could be made (Bisdikian et al., 2014). For example, a standing policy by stakeholders might be to obfuscate or suitably alter measurements from their sensors, before sharing them with other members.

A decision-maker operating in this kind of environment can benefit from our model to address the issues outlined. By forming groups of similar sources, redundant sampling can be minimised. This could potentially prolong the life-span of a sensor network. A learned sampling strategy would enable the system to adapt to different task requirements. For instance, more sensors could be sampled in critical situations, where the benefits of making accurate assessments of the world outweigh any potential costs (e.g., in cases of potential threats from terrorists). Also, source diversification and strategic sampling may entail that sensors or sources from a less friendly nation is avoided or sampled less in relation to sources from allies. Our diversity model is particularly crucial in this application scenario, since sensors owned by the same stakeholder may be expected to provide information that aligns with the objectives of their owner or host nation. Therefore, regarding them as independent sources may only work to the detriment of a decision-maker.

6.2.3 Diversity and Argumentation

When considering multiple sources espousing multiple claims, it is possible that these claims are presented by way of more logical arguments in some applications contexts. In such environments, for instance, a decision-maker may engage in dialogue with sources, exchanging arguments, and probing these sources for additional information (or clarifications) concerning their claims about the state of the environment (Oren et al., 2007). Also, sources, in addition to their reports, may provide insight into the rationale behind the claims they have made; the reason why or how they arrived at such conclusions, their thought process in general, which could strengthen or weaken their contributions (Stranders, 2006). This kind of system provides a more elaborate mechanism for dealing with conflicts when integrating multiple contributions (Şensoy et al., 2013; Villalta et al., 2011).

We believe that logical argumentation frameworks (Bench-Capon, 2003) can add some interesting dimensions to our current model and to the truth discovery process in general. For instance, it would be interesting to explore how these schemes affect the process of group formation. Would it be reasonable to place in the same group sources, who might disagree in their reports, but the rationale behind their claims tend to agree in some ways? For example, a group of sources belonging to the same organisation may ordinarily provide similar reports. However, in some cases, the process of transmitting those claims may be compromised way beyond their control, thereby introducing inconsistencies. In such cases, the arguments or rationale behind their claims could still serve as evidence for deciding how to classify the sources. Earlier we reflected on the possibility of having a reputation overlay to enable the sharing of possible information that would assist decision-makers in creating diversity structures. Such provision, if made available, could also benefit from argumentation techniques. For instance, it would enable agents to argue about the suitability of different similarity metrics or feature relevance in different contexts. A similarity
metric that is claimed to have worked in a similar setting to that currently faced by a requesting agent, might be more favourable and likely to be adopted than those that do not match the considered settings.

6.2.4 Richer Features Representation

One other way we feel our current model could be enhanced is in the representation of features. In real-world contexts, there may be complex relationships between features. For example, a source working for a subsidiary company may still be bound by the organisational ethos of the parent company. In this regard, sources from these two related companies might actually be considered similar. Our current feature representation does not allow us to capture these sort of relationships. Such domain-specific knowledge can be represented as ontologies, and these relationships if they exist, could be exploited in forming better groups. A similar mechanism has been exploited by Şensoy et al. (2014) in their formulation of stereotypes. Not exploiting this kind of knowledge may actually lead to degrees of diversity that are not actually necessary, such as forming more groups based on the assumption that features are orthogonal.

6.2.5 Working with Data Streams

Our current work uses a tree induction algorithm that relies on static information, and therefore operates on the assumption that the entire training set is available. However, if the input is a continuous stream of data, a batch algorithm such as classical decision trees or model trees may not be tractable as new information/evidence continues to arrive. Processing delay is an issue, and may not be acceptable in some domains. Some form of incremental learning algorithm that scales linearly with the incoming data may be beneficial. A number of authors (Potts and Sammut, 2005; Ikonomovska et al., 2011; Gama et al., 2003) have looked into the problem of decision tree induction in contexts of high-speed data streams. This usually involves some form of incremental tree induction, where sufficient statistics is maintained at each decision node, and decisions to split are made only when there is enough evidence in favour of a particular split test (Gama et al., 2003).

A similar interesting research opportunity presents itself in attempting to estimate environmental states using incoming report streams (Etuk et al., 2013a). This is a non-trivial scenario involving timely assessments of trust in fused information under a rapidly changing world, and may require several decision points, for instance, to determine appropriate stopping criterion, where a decision-maker ceases to process incoming reports and inferring on a proposition.

6.3 Summary

In this chapter, we have reflected on our research, discussing the limitations of our model as well as potential applications and avenues for future research. In particular, we have developed a model based on the assumption that correlations exist between features and reports of sources. Although available evidence suggests that different kinds of dependencies exist among sources in many environments, the applicability of our model may be restricted in contexts where dependencies among sources do not correlate with their features.

Our decision-making model is based on reinforcement learning that is computationally expensive. Although the formation of groups is beneficial in managing this complexity, as the size
and number of groups increase, the computational cost of learning sampling strategies may increase significantly restricting the applicability of the model.
Chapter 7

Conclusions

The focus of this research has been on developing effective mechanisms for truth discovery in resource-constrained environments. The primary question we have attempted to answer is: how can we strategically sample information sources for opinions regarding the state of the environment to acquire as accurate an estimate as possible, given that there are costs associated with acquiring information and we need to operate within resource constraints?

In summary, the contributions of this research are:

- **A general framework for multi-source integration under resource constraints**: we presented a framework, TIDY, for truth discovery when integrating evidence from variously trusted information sources under resource constraints. By optimally sampling the source population for evidence while maintaining the quality of information, TIDY significantly increases the value of information acquired from sources for a decision-maker.

- **A mechanism for source diversification**: we presented a mechanism for source diversification that minimises redundant sampling and also mitigates the effect of correlated biases. This mechanism also offers a decision-maker the opportunity to exploit evidence from groups with potentially different perspectives on a problem. The model of diversity was evaluated in the context of a simulated resource-constrained multi-source environment, and shown to perform well in estimating ground truth. In particular, with an increasing proportion of dependent sources, an approach based on diversity performs significantly better than approaches that are not based on diversity in making accurate estimates of environmental states. When the assumption of dependency is relaxed, the model does not perform any worse than approaches that do not take diversity into consideration.

- **A decision model based on diversity and trust**: we presented a decision-theoretic approach that combines diversity and trust to identify effective sampling strategies. This model allows a decision-maker to balance, in a principled manner, the trade-off between quality and costs of acquiring information. Again, our model was evaluated within a simulated environment, and shown to be effective in guiding sampling decisions. In particular, the results of our experiments demonstrate that our model yields higher utilities than approaches based only on diversity or trust.

We believe that the value of this research has far-reaching impact especially as modern society continues to be bombarded by myriads of information sources that can support decision-making.
CONCLUSIONS

By developing efficient methods of harnessing their potential, we believe that we have contributed to addressing the existing problem of too much information but not enough knowledge. While we have described our model in general terms, we have been careful to reflect on some concrete examples and application settings where our work can be applied. In particular, we believe that the problem we have addressed in this research has significant impact in environments such as sensor networks where working within resource constraints is critical, and in social networks and many crowdsourcing problems where dependencies among information sources increase the risks of biases.
Bibliography


Bush, G. (2003). The national strategy for the physical protection of critical infrastructures and


Fusion, pages 1–7.


Surowiecki, J. (2004). The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations. Random House, Inc.


