LITERATURE SURVEY

S.NO	TOPIC	AUTHOR	YEAR	ALGORITHM	ACCURACY
1	Temporal Summaries of News Topics	James Allan, Rahul Gupta, and Vikas Khandelwal	2001	The surprising result for usefulness is that a round robin ranking algorithm performs almost as well as useful2. We believe that reﬂects the pyramid nature of news reporting: important, and therefore probably on-topic, information is reported early in a story. Later material is	70%

				more likely to be tangentially related to the topic, and so ranking it lower helps	
2	Introduction to information retrieval	Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schu¨tze	2008	n the context of Information Retrieval, link analysis focuses on ''the analysis of hyperlinks and graph structure of the web'' with the intention of facilitating web search. In particular, chapter 21 provides quite a comprehensive introduction to Google's PageRank algorithm	70%

3	Topic Detection and Tracking	James Allan	2002	TDT tasks, Story Segmentation, First Story Detection, Cluster Detection, Story Link Detection	73%
4	Breaking News Detection and Tracking in Twitter	Swit Phuvipadawat, Tsuyoshi Murata	2010	To ensure that messages in the groups are related to the first story and to allow further messages to develop upon previous messages we will compare a message with the first message in a group and the top k terms in that group. In this experiment we set k to 10. We then assign a message to a group if the score exceeds a pre-defined threshold	78%

				called MergeThreshold. The results to this point are groups of messages or news stories.	
5	Tracking of Data Leaving the Cloud	Yu Shyang Tan*, Ryan K L Ko*, Peter Jagadpramana*, Chun Hui Suen*, Markus Kirchberg*, Teck Hooi Lim*, Bu Sung Lee*, Anurag Singla†, Ken Mermoud†, Doron Keller†, Ha Duc	2012	is to send these log files back to the server for further processing. This is done at the end of each user session. When a user session is closed, the viewer program will attempt to establish a connection with the main server back in the cloud domain. If successful, the viewer program will encrypt the files so as to ensure the integrity of the logs and sends the cipher	65%

				back to the server. Standard AES and RSA encryption algorithms are used in the encryption process. In the event where the connection with the main server cannot be established, e.g. no internet connection, the viewer will delay sending the log files until the end of the next user session	