Summarization 1

The battle of Arawe was fought between Japanese and Allied forces as a part of the New Britain (the largest island of what is today Papua New Guinea) Campaign in the Pacific theater during World War II. Rabaul, the capital of New Britain, was a major base for the Japanese military in the area, and in 1943, the Allies initiated an offensive to contain the numerous Japanese forces there. Although the offensive on Arawe was largely a diversion from other parts of Operation Cartwheel, the Allies secured Arawe after roughly a month of fighting. It has been questioned whether the battle was really necessary, or if the Allied troops could have been of more use elsewhere.

Summarization 2

The Battle of Arawe (also known as Operation Director) was fought between Allied and Japanese forces during the New Britain Campaign of World War II. The battle formed part of the Allied Operation Cartwheel, and had the objective of serving as a diversion before a larger landing at Cape Gloucester in late December 1943. The Japanese military was expecting an Allied offensive in western New Britain, and was reinforcing the region at the time of the Allied landing in the Arawe area on 15 December 1943. The Allies secured Arawe after about a month of intermittent fighting with the outnumbered Japanese force. In July 1942, the U.S. Joint Chiefs of Staff directed that the main objective of the Allied forces in the South Pacific and Southwest Pacific area commands was to capture the major Japanese base at Rabaul on the eastern tip of New Britain.

Summarization 3

Initial Allied goals for the landing at Arawe included securing a base for American PT boats and diverting Japanese forces away from Cape Gloucester. The Japanese air units at Rabaul had been greatly weakened in the months prior to the landing at Arawe by prolonged Allied attacks and the transfer of the 7th Air Division to western New Guinea. This is an order of battle listing the Allied and Japanese forces involved in the Battle of Arawe from 15 December 1943 to 24 February 1944. It was intended that PT boats operating from Arawe would disrupt Japanese barge traffic along the southern shore of New Britain and protect the Allied naval forces at Cape Gloucester from attack. After organizing his force while waiting for the 1st Battalion, 141st Infantry Regiment, Komori began his advance on Arawe on 24 December.

*Two of these were generated by computers,*

*one of them was written by a human,*

*can you tell which of them is which?*

Naïve approach

To generate a summary of N sentences, we pick the first sentence from the first N paragraphs.

That’s it.

Worth noting is that this algorithm is extremely fast, and consistently produced high subjective readability.

**AVERAGE RESULTS**

|  |  |
| --- | --- |
| **ROUGE-1** | 0.59 |
| **ROUGE-2** | 0.20 |
| **ROUGE-3** | 0.06 |
| **ROUGE-4** | 0.02 |

LexRank

LexRank works a lot like PageRank. It was found that it did not perform well in summarizing Wikipedia articles.

Each sentence is represented as a vertex in a graph. An undirected edge is created between two sentences if they are similar enough, ie above a certain threshold for TF-IDF cosine similarity.

Power iteration or the Monte Carlo methods are then used to obtain the LexRank score for each sentence.

The summary is generated by choosing the top N ranked sentences according to their LexRank score.

LexRank can be used to summarize one, or a cluster of documents.

We filtered out duplicate sentences in the cluster, as well as short sentences (less than nine words) as a pre-processing step.

**AVERAGE RESULTS**

|  |  |
| --- | --- |
| **ROUGE-1** | 0.35 |
| **ROUGE-2** | 0.06 |
| **ROUGE-3** | 0.00 |
| **ROUGE-4** | 0.00 |

LexRank + Heuristic Rules

Similarity to query

Cosine similarity between query and each sentence.

Positional score

Sentence with index *i* in section with index *j* in article with index *k* in document collection *A*.

**AVERAGE RESULTS**

|  |  |
| --- | --- |
| **ROUGE-1** | 0.57 |
| **ROUGE-2** | 0.24 |
| **ROUGE-3** | 0.09 |
| **ROUGE-4** | 0.04 |

**EVALUATION METRIC**

**ROUGE = R**ecall-**O**riented **U**nderstudy for **G**isting **E**valuation

To score the generated summaries the ROUGE-N method was used. The method measures the recall of N-grams compared to a human written summary of the same text.