

White Paper on CallMiner's Tireless Supervisor™

Executive Summary

CallMiner's Tireless Supervisor product is capable of taking poor fidelity, multiple speaker recorded audio and accurately identifying and reporting on key topics and concepts, uncovering caller intent, classifying calls by subject, and automatically scoring calls defined by the user.

Tireless Supervisor accomplishes this by using patented technology that identifies patterns of speech that quickly report on the content of a call without having to manually listen to the recorded call. Since millions of calls are recorded hourly, it's been impossible up to now to divine the intelligence buried in recorded calls. Now this is possible, practical, and paying huge dividends in real information.

Problem Statement

The volume of recorded calls makes manual listening to the calls, to gather intelligence, uneconomical. Call centers would have to more than double the number of employees to listen to all customer/agent calls, and still not have enough information to analyze the results. The majority of contact that businesses have with customers is through the call center, and this valuable source of business intelligence has all but been lost to the archives of backup tape. The current process is typically to randomly select a few calls. While this may or may not give a statistically valid sample that can quantify overall results, it can't go beyond that. It gives information, but little insight. An automated process that can score all, or at least a significant portion, of the calls can achieve the current results automatically and eliminate the manual resources. Or, more interestingly, the automated process can refocus those scarce resources to the calls that need attention.

Over the years, speech recognition technology promised to deliver a means of digging into this vast corpus of recordings and unearth what the callers were saying, but unfortunately, analytical tools to draw conclusions from this information were lacking. Speech recognition technology; specifically speech to text applications, has been around for over 20 years and for most applications does not provide a level of accuracy that yields significant intelligence. A very good, speaker trained, high quality microphone dictation system still is only about 90% accurate. When you apply these technologies to call center audio for purposes of mining, this accuracy number plummets. The accuracy rate is further exacerbated by multiple speakers talking over each other, ambient noise captured with telephone based recordings, non-directed and non-trained conversational speech, and finally, compression in the

stored call from which most speech recognizers are lucky to get a 50-60% accuracy rating.

Add to this low recognition rate the desire to go well beyond just finding a few words in a conversation, but instead attempting to derive meaning, understand intent and context of a conversation. The end goal being able to correctly classify and score each call that is recorded, accurately, automatically and quickly.

Goal

To take the low speech recognition rate and attempt to derive meaning, understand intent and context of a conversation. Then correctly classify and score each call that is recorded, accurately, automatically and quickly.

CallMiner Solution

CallMiner developed patent pending technology to find data, gather intent, classify and score calls, using the output of large vocabulary speech recognition technology that was specifically developed by CallMiner. Tireless Supervisor™ uses the idea of intent pattern recognition (IPR) and pulls meaning out of the seemingly useless results generated by speech recognition. The user defines the scoring process, allowing the system to be focused on identifying calls of interest for any reason. It may be particularly 'good calls' (branding, courtesy, cross selling) to provide insight on effective techniques. It may be focused on 'bad calls' to identify problems. Currently, the random process means that the manual listener hears a very high percentage of uninteresting calls and a small percentage of interesting calls. Tireless Supervisor scores the calls automatically which allows the manual listener to then focus on interesting calls.

Understanding the Problem space size

The CallMiner CoreEngine produces an average of 8-10 data points for each word spoken during a typical conversation. This same typical call center conversation is about 400 words in length, meaning CallMiner produces a potential speech matrix of 4000 data elements. Tireless Supervisor's IPR engine uses a combination of fuzzy pattern matching and advanced permutation analytics to evaluate the entire speech matrix. Consider that the speech matrix has a potential size of

$$8^{400} = 1.7218479456385750618067377696053e+361$$

Through special algorithmic pruning, Tireless Supervisor typically has to evaluate only several hundred thousand.

Finding the Intent

IPR relies on the concept that all conversational threads have, at their core, one or more “patterns” of speech. These patterns of speech don’t necessarily have to contain exact word matches for particular search terms; rather they have to “look” like the specific pattern. Each pattern is assigned a weight by the pattern developer, and each area of intent is assigned a threshold. If a group of patterns match, and their added weights exceed the threshold, then that conversation is said to “look” like and contain that intent. The weights and threshold are user definable and easily tweaked to produce better and more accurate results.

Sharing the Patterns

A typical intent “bucket” will contain anywhere from 5 to 100 “patterns” to match. The good news is that these patterns tend to apply across domains, meaning that a pattern block for matching a credit card transaction is valid in the banking industry and in the travel industry. So as the pattern base grows over time, all users of this system can benefit from more accurate and more finely tuned patterns.

Empirical Results

Does it work? As stated previously, the average call center audio speech recognition rates are in the 50%-60% range, while the Tireless Supervisor system has produced Business Accuracy rates of 95%-98%, by correctly identifying what is being said in the call. Business Accuracy is the ability to correctly identify the intent, content and classification of a call, including the ability to assign multiple areas of intent and classification.

Conclusions

CallMiner’s Tireless Supervisor product overcomes the inherent accuracy and technological shortcomings of standard speech recognition in such a marked way that business can finally extract real value from recorded conversations. By focusing on intent and content of the call rather than trying to exactly match what is said, and going beyond just simply word spotting technologies, business can for the first time really gain insight into what their customers are saying.