100 MILLION PROGRAMS IN 3 CHARACTERS: INNOVATIONS IN SEARCH TECHNOLOGY FOR MASSIVE DATASPACES

Rakesh Barve, Ph.D.
Veveo, Inc.

Abstract

The sheer volume of video a consumer can watch on a TV in the comfort of the living room is becoming overwhelming; unfortunately, user interface design has not kept pace with the UI advancements and innovation we have seen in desktop computing. In addition, TV suffers (as mobile platforms do), from the constraints of restricted screen real estate and sub-optimal input mechanisms (on-screen keymats, directional keys, number keypads, qwerty keypads, etc.). This combination of a dynamically increasing amount of content and sub-par browsing paradigms has accentuated the need for better and more powerful search functionality for the TV experience.

Building a better search solution for the living room TV experience poses a number of challenges for the TV aggregator:

- *Massive growth of programming choices*
- *Multiple data sources to be indexed and merged for one search box/results list*
- *Metadata enhancement & consistency*
- *Easy Input – search-as-you-type, incremental results*

In the past, MSOs have struggled to improve their search capabilities and have accepted limited search functionality, relying on basic techniques such as purely lexical-based matches, (as opposed to relevance-base techniques similar to how Google works). However, most aggregators recognize the value of search and the need for improvement.

There are numerous technical challenges and user experience issues involved in building a better search solution. In this paper we propose a novel search and browse system that is based on the principle of minimizing a user's input, (i.e. the number of keystrokes), to get to the desired content. We outline how the system needs to simultaneously support maximal-recall, field agnostic querying, and incorporate enough of the user's application context that the number of input key-strokes is minimized (by squeezing as much information as possible from each key-stroke).

CONTENT DISCOVERY

Search And Discovery Components

vTap's basic *Incremental Search* paradigm helps users to easily find specific individual videos from a plethora of videos. vTap's *Universal Smart Tag Clustering Engine* creates a huge number of independently searchable collections of videos corresponding to every known person, movie, TV-show, music band, sports team, and 'micro-genre' in the world, by analyzing the text meta-content, (the subject, genres, fields and other attributes), of individual items. Each searchable collection is like a channel; there are a virtually unlimited number of semantic, subject-specific channels. Based on the nature of the

collection and current, (TV, VOD and internet), availability of pertinent items, these listings may dynamically evolve over time when the user browses them. Furthermore, vTap's *Contextual Search* allows easy search and dynamic creation of *playlists* of individual items within the limited context of a specific named collection. When deployed, this feature enhances the user's ability to browse into a particular collection's listing and then refine it further while still in that context, creating a playlist on-the-fly. In addition, vTap also uses a *Learning Engine* that observes the Smart Tags, micro-genres and keywords associated with videos watched by a user, and implicitly infers a stochastic video signature of that user that captures his taste in video, (the smart tags and micro-genres), along with TV/VOD related habit patterns and subscriptions. A *Recommendations Sub-System* then uses the stochastic signature of the user, (inferred by the Learning Engine), and smart tags therein, as well as explicit smart tags set by the user to recommend relevant individual videos encountered in the server's dynamically evolving TV/VOD/streaming video database. A *Usability Personalization Sub-System* uses the stochastic signature, (the smart tags dimensions as well as specific habit related information), to make possible personalized reordering in Search, Browse and Actionable listings that improve usability in significant ways.

Incremental Search – A New Search Paradigm

The way vTap Search enables users to easily find content is via the network-based *incremental search* technique – for each character input by the user, (without hitting a 'return' or 'enter' button), a network search operation instantaneously returns a new set of corresponding results that enable the user to 'converge' to the desired result with minimal text input. vTap's incremental search is fundamentally different from other approaches, which incrementally help users complete search queries by performing a strict prefix-based lookup on a database in which each record is simply a query made in the past. In contrast to these approaches which return completed *query candidates,* one of which needs to be picked up by the user to launch the real search that returns results, vTap Search directly returns *results* which can be meaningfully acted upon.

In vTap Search, each record is not a simple string, but is in fact a composite structure comprising multiple text fields – for example, title, tags, descriptions, source-sites etc. Specifically, vTap incremental search is characterized by the following attributes:

- On any single-word or multi-word query, it simultaneously checks across all fields for prefix-matches with any of the query words.
- Not only are multiple query-words in user's incremental input string allowed to be incomplete, these incomplete query words can each prefix-match a distinct word in a distinct field of a composite record.
- Having multiple fields inside a record allows for calculating more detailed match-scores, based on the matching fields, which in turn helps to better rank results. Field match scores can be based on the semantics or importance of the fields from a search-recall perspective.
- Features such as giving results even when the query has spelling errors.

Incremental Search Intellectual Property

One way to implement the incremental search just described is to simply scan every word in every meta-content record of the video database and then display matching

results. When the database of records is small, (e.g., hundreds of names in a contacts database), naïve techniques will suffice. However, clearly this approach of going through every word of every record in the video database for every query does not scale, (say, for tens of thousands of movies/music artists/cast, millions of music tracks and internet videos etc.). The fundamental innovation in vTap Search is the mechanism by which only a small fraction of items need to be looked up in order to return meaningful results for incomplete queries. All of the following work in unison as part of vTap incremental Search:

(a) the video and cluster/collection popularity computing algorithms for different kinds and types of videos, (vTap uses a variety of heuristics and internet sources to assign popularities as described in the next subsection),
(b) the pre-processing, static indexing steps and the term relevance logic which blend video and cluster popularity with the relative importance of the meta-content fields that contain the term,
(c) the custom runtime incremental search data structures,
(d) the runtime query processing and term relevance re-calculation based on how and which meta-content fields matched the query terms
(e) runtime relevance changes based on scheduled time, subscriptions--availability-dependent, (e.g. if a program VOD window just expired), TV schedule attributes such as LIVE or reruns etc., and user-signature dependent dynamic runtime relevance changes and,
(f) the custom runtime computation steps that bring it all together.

vTap Video and Cluster Popularity

vTap Search applies to text based search for items that may be TV, VOD or internet videos or named smart tag clusters of those videos.

The vTap service's video database of available TV and VOD videos is updated regularly using TV and VOD guide feeds available from various data suppliers. In addition vTap uses a variety of other internet sources to secure richer information on every conceivable movie and TV program. Internet link, modified page rank and rating based heuristics are used to assign popularities to movies and TV programs from various internet sites first. Movies and programs available in VOD and TV listings of the vTap video database are correlated to corresponding entries in a universal database of all movies and TV shows, and these listings then derive a base popularity metric from the corresponding entries of the universal database. They also derive enhanced text meta-content, (more terms for indexing), based on the correlated entries. TV show popularity can be further adjusted by taking into account when the show is telecast, the nature and popularity of the network telecasting it, whether it is a current TV show or not, the cast in the show and its popularity, and so on. Similar rules can be applied to VOD content, except there is no telecast time there -- although there are notions of "recently available."

There are different kinds of internet videos based on their source. Some videos – e.g., news, sports and those from content companies, have a dominant recency component in their popularity computation and are also topical in nature based on meta-content and the popularity and authority of the site producer. Videos from other content companies depend roughly on how well

known or popular the consumer facing content website is. While calculating the popularity of user-contributed internet video, (since most videos do not have external referring links or very few links), the notion of internet page rank or popularity is of limited use. So for user-contributed videos, one must make use of 'social graph' information, internal page ranks among the site's uploading users and subscribing users in various ways, and statistics such as view counts, number of comments and recency. Many popular video sharing sites have content partners uploading videos regularly, and the popularity of these videos is again a function of partner popularity and recency.

The popularity or relevance of specific collections, (Smart Tags), of videos is a function of the corresponding topic's 'popularity' overall and in its domain, and also the temporal properties of the videos associated with that topic.

Term Relevance For Each Term In The Text Meta-content Of A Video

The vTap paradigm of incremental search makes it very important to assign appropriate 'search' weights to appropriate keywords, terms and meta-content in the specific subfields associated with a video. Text meta-content for TV and VOD listings have fields such as title, episode title, keywords and crew members -- correlation with internet sources of movies and TV shows can enrich this meta-content. Internet videos typically have titles, descriptions and tags. User generated videos have very basic meta-content, whereas videos from content companies have editorially created richer meta-content.

Web search engines use information and signals embedded in the keywords and associated with various html tags and page structure at the time of indexing. When indexing videos from TV, VOD and the internet, Veveo can use the fact that text in different fields, (title, description etc.), is of varying importance. The basic advantage of creating a structured database of multi-field records, is that it allows the same record to be recalled using multiple associated names and keywords with different search weights: for example, also-known-as names and, (where appropriate), abbreviations can also be used to index the same record. Additionally, some of the fields -- e.g., running time, video format, dates etc., allow for a filtered or sorted view of search results.

vTap Search 'Number Mode'

When using vTap incremental search from devices such as PDAs that have a QWERTY keyboard, the user's input is unambiguous as each keystroke is mapped to an individual character. However on most phones and on all conventional TV remote controls, there is typically only a numeric 'phone' keypad and inputting a single character using these devices may frequently require pressing the same key two, three or four times. The associated inter-digit timers make the process of inputting a query to a search engine from numeric keypads both error prone and cumbersome. Contrast this with the relative ease of composing SMS or text messages in the so-called 'predictive-text mode.' This ease is a result of the fact that more often than not, the user presses a single keystroke for each character of each word in the composed text message. However, as far as the search application goes, assuming the database being searched is a huge web-scale database, even the 'predictive-text mode' technology used to create text messages from such phones is completely useless in practice. This is because the device resident 'predictive-text mode' technology is premised on the practical assumption that there is a *very restricted list of words* that constitute the 'working set' of words a user

will draw from in order to create text messages. In stark contrast, the set of terms that a user has at his fingertips to launch a search query is a *mammoth set*, comprised of all kinds of noun phrases, all (full, short, nick) names and proper nouns of every entity, (including movies, TV-shows, bands, songs etc), or persons that have been named, 'user ID's of various kinds, words and terms from all languages, including multi-word noun phrases and so on. This renders the 'predictive-text' approach completely unworkable in a web-scale search application.

In order to *dramatically ease* the experience of using vTap incremental search from such numeric keypads, vTap provides a 'number mode' -- based on Veveo's proprietary technology geared towards the incremental search principle -- wherein the user needs only to press a single key, (the numeric digit corresponding to the character), for each character. The ability to input queries quickly, especially without having to pay any heed to inter-digit timers, and then receive results in the vTap incremental search style is what dramatically enhances the search experience from phones and TV remote controls. Even though the user's query is inherently ambiguous, ('227' could mean 'car,' 'bar,' etc.), vTap computes and blends results corresponding to all meanings of the multiple input query tokens according to highest relevance. Even in the number mode, vTap uses all aspects of its incremental search algorithm so that *all of the features mentioned work in the number mode,* including auto-corrected spelling matches. vTap displays matched strings to clearly enable the user to recognize the records of interest among the results.

A Comparison Of vTap's Character-based Incremental Search vs. 'Standard' Word-Based Commercial Search Engine Techniques:

Since the results change with each keystroke input by the user, the user can get to the result of interest with a minimal number of keystrokes, (the query words are allowed to be prefixes of words in the result). This creates a tremendous advantage over the cumbersome character input on devices like phones, TVs and Mobile Internet Devices.

The technology and algorithms used to implement vTap's incremental search are distinct from the ones used in standard word-based search engines. In the latter, the user types one or more words and hits an 'enter' button so that the user's input step and results step consist of *discrete distinct phases*. Because the query almost always contains full words, it can essentially look up pre-computed lists of matching records, (one list for each word), and finds the best records amongst those that are present in all or most of the lists. Moreover, the items that are indexed are not records with fields but unstructured html web pages.

On the other hand, in vTap, the input and results processes are finely interleaved and interactive, so that the user may get to the desired results with a minimal number of input keystrokes. So, at every keystroke, an implicit query is launched to the server and in general, the query may consist of a space-separated list of multiple incomplete prefixes. Using the same approach as the one described above for word-boundary based search would require the pre-computation of one distinct list of records for each prefix of each term in the system. This is infeasible even on huge infrastructures when the number of indexed terms and records in the system is large enough. In order to provide

the incremental effect, vTap's search technology uses a unique set of search data structures and algorithms in a special purpose information retrieval system focusing on instantaneously returning relevant results in response to each keystroke input by the user. The key aspect that enables this system to very quickly compute relevant results for each keystroke even when the search database is massive, is the way it combines offline relevance computation, search data structures, and runtime algorithms; in that way, the result computation process has to examine only a small fraction of the entire search index.

vTap's Universal Smart Tag Clustering Engine

vTap's Clustering Engine provides a practically unlimited number of specific collections of videos which are the rough analogue of conventional 'TV channels.' In order to facilitate very diverse *and* specific interests, vTap provides not just the usual broad categories of TV/VOD/ streaming videos, (news, sports, etc.), but also a huge number of specific and *dynamically evolving* collections of video items. Each collection corresponds to a vTap Smart Tag. Basically, every named person, movie, TV-show, music artist, sports team or entity is tracked by vTap in a Universal Smart Tag database. As and when the system discovers more smart tags, these are added to the UST database. The main motivation is that in the entertainment domain, the user often wants to get to a specific set of items or a topic and then browse through the items all at once, rather than always performing a search for each single item. Furthermore, the UST based clustering forms the basis of capturing the video taste and video interests in the *Learning Engine,* which is used in the *Recommendations sub-system, Personalized Usability sub-system* as well as in analytics. In vTap Search, named collections, (each collection corresponds to a smart tag), and single items are appropriately blended in the search results, so that each collection can be independently retrieved using incremental search based on the associated meta-content.

The clustering engine works by analyzing the meta-content crawled and mined for each video or clip, and then using clustering techniques to associate each video with one or more named specific smart tags, or broader 'genre' categories. Examples of specific smart tags include George Bush, the Boston Red Sox, Tom Cruise, Jay Leno, Herbert Von Karajan, Richard Feynman, Otto Von Bismarck etc. In short, the clustering engine creates specific collections representing meaningful Smart Tags from domains such as news, science, music, sports, movies, TV, as well as topics based on personalities and other named entities.

As a result of the dynamic nature of the TV schedule, VOD asset availabilities and changing websites, the Smart Tags are also dynamically evolving: new smart tags are added as part of the TV/VOD availability changes, crawling, discovery and clustering processes. The relevance of the collections among search results also varies: video items belonging to any given collection may change based on new videos that get discovered and fall into that collection, old videos no longer available on TV or VOD or whose web links become stale have to be removed, and the ordering of videos themselves may be optionally changed when the user decides to browse a collection. Examples include: news oriented topics listing the latest videos on top; sports topics supporting a dichotomy between videos corresponding to scheduled LIVE games, videos featuring highlights of the games themselves, (as opposed to more generic news videos on the same); movie topic previews and interviews, as well as VOD and TV availabilities of full movies and so on.

In vTap, the default organization of videos of a collection may depend on the topic of the collection as described above, and be segmented into TV, VOD and internet video sections. However, vTap can also allow the user to sort and browse the videos in each section as appropriate.

vTap's Contextual Search And User-Specified Dynamic Playlists

As previously mentioned, named collections, (i.e. Smart Tags), of related videos are included as individual search results to a user query. This is useful especially when the user's intent is tentative, which is more typical when the user is in the mood for entertainment as opposed to conducting a typical internet search. For instance, the user may start off with the intent to watch Jay Leno videos in general, but after watching one or two, decide to watch specific Jay Leno videos lampooning George Bush.

Using vTap, a user can quickly get to the Jay Leno video collection within a few characters and then begin to peruse the list of Jay Leno videos in the collection. After watching one or two, vTap allows the user to very intuitively conduct a *contextual incremental search* within the collection by typing in more characters. So, for instance, the user could start typing in 'Bush' and vTap would then pull up Jay Leno videos that mention George Bush in their meta-content. Going further, the user can easily play this list of George Bush videos one after another as a *video playlist*. Effectively, this is a user-specified dynamic playlist created by first tentatively focusing on the full subset of Jay Leno videos, and then on-the-fly doing further incremental search to refine the playlist. Enabling this user experience and fine grained user-control is a capability unique to vTap Search.

The playlist aspect of this feature is more relevant to internet videos that can be instantaneously streamed, but contextual search makes sense even for TV and VOD -- for instance, a user may enter the collection of available Seinfeld episodes, browse through them, and then continue to narrow down to one episode by making a search query restricted to that collection.

vTap Learning Engine

vTap's learning engine analyzes, for each user, the history of videos and collections that have been viewed, the search history, the smart tags, micro-genres and keywords associated with the viewing history, and the preferred TV channels, VOD subscription packages or web sites corresponding to the videos. From this information, it creates a 'stochastic signature' that can be said to represent the video viewing habits of the user. In its base case, the stochastic signature captures the smart tags, (personalities, TV shows, music artists, sports favorites etc.), micro-genres and TV channels, VOD subscriptions and websites that the user has a tendency to view when it comes to video. In a more advanced deployment, the stochastic signature can even capture the time of day and day of week that the user typically views videos associated with each smart tag or TV channel etc. The Learning Engine architecture is general enough that given adequate information, the stochastic signature can capture specific TV/VOD viewing habits of the user such as whether he watches HD or SD, paid or free VOD, and other habits that the service provider deems worthwhile.

The stochastic signatures across all the users can potentially also be leveraged to answer analytics related queries for the Service Provider -- for instance, what is the correlation between users who watch paid VOD movies and users who watch subscription sports channels on TV?

vTap Recommendations Sub-System

vTap Recommendations for discovery have two distinct flavors, explicit and implicit. In 'feed'-based recommendations, the vTap backend system allows users to explicitly specify a list of smart tags they are interested in, and whenever vTap comes across any new TV, VOD or internet video in its backend database that is associated with the user's smart tags, those videos are sent across among the Recommendations. This allows vTap to flag user videos pertaining to the user's explicit interests, and enhances the probability that the recommended videos will be watched.

vTap's Recommendations sub-system also uncovers information that is more implicit in nature, such as information about unviewed videos that share the user's stochastic signature characteristics, and are available in the user's subscribed TV channels, VOD package or somewhere on the internet. In addition, the Recommendation sub-system can employ collaborative filtering techniques to cluster 'similar' programs/videos, (or 'similar' users), based on their viewing habits. The Recommendation sub-system can also recommend to one user other interesting videos 'similar' to the videos or TV shows from his stochastic signature, or similar to the stochastic signature of other 'similar' users.

vTap Usability Personalization Sub-System

The Usability Personalization Sub-system is simply a system meant to apply the stochastic signature, (computed by the Learning Engine), to personalize and reorder various Search, Browse and Action listings in various contexts.

For instance, the stochastic signature tracks the list of TV channels, (part of the smart tags), as well as the list of TV shows, crew etc. derived from the user's viewing history. This means that to present a user with a list of (implicitly computed) Favorite Channels or Favorite TV Shows etc., all that is required is a lookup in the stochastic signature. On the other hand, a personalized reordering of the Movie channel listing in the TV guide application, is a matter of taking the original list of Movie channels and then hoisting above other channels all the channels that appear in the user's stochastic signature. Personalization for search re-ordering, (e.g. search followed by a one-click action to tune to a channel), enables the particular user who simply types a 'c' to bring up 'The Church Channel' before the globally more popular CNN. This is achieved by taking the first few pages of search results and ensuring that results corresponding to the stochastic signature are hoisted up. However, each reordering needs to be done carefully, to avoid hoisting search results that have low relevance terms matching the query.

If the stochastic signature is rich enough to capture the HD/SD preference of the user, then whenever a channel or program appears on multiple tuner numbers, the default one-click tuning action can be made to resolve to the tuner that matches the user's HD/SD preference. If the stochastic signature is rich enough to capture attributes such as free versus paid VOD content, it can be used to reorder VOD movies browse listings to match the user's preference. In general, several other usability personalizations can be achieved by ensuring the stochastic signature captures the relevant preferences, and by looking up the stochastic signature at the time the user is actually searching, browsing or about to select an action.

INTELLIGENT AND SCALABLE WEB VIDEO CRAWLING

Veveo crawls videos from thousands of different websites and presents these in its search play list results. Currently, vTap's

growing video index consists of at least 250 million internet videos. Crawling the web for videos is fundamentally different from crawling the web for web pages for the following reasons:

- In the web search application, when the user types any set of words, the entire web page is the result. In video search, only the specific video is the intended result, and this is tedious when the page contains multiple videos and the crawling and information extraction system must ensure that the meta-content corresponding to one video is associated only with that video and so on.
- Moreover, the crawling system must be able to semantically interpret specific strings in the meta-content, unlike in traditional crawlers. Therefore 'today' in the date field must be converted to the date value on the day of crawling: similarly, clip attributes, e.g. tags, various clip dates, user-IDs, descriptions, genres, and various statistics need to be extracted with a deeper understanding of the content layout and offering, sometimes specific to the site.
- Very often, the crawling system involves the ability to analyze scripts/programs to automatically detect the presence and play-attributes of video links.
- The video link that one comes across could be permanent or transient -- again a deeper understanding and automated analysis of the site is required to ensure the stored URL is permanent and not transient.
- Re-crawl systems have to be optimally and specially designed to ensure that the 'liveness' or 'staleness' of a video URL is known correctly.
- A re-crawl process design is also necessary to track various statistics associated with the video and other videos linked from and to it.
- Where available, the nature and number of uploading users is information that is crawled and stored.
- When there is a change in the crawled page's format, there has to be a real-time change in the crawler to adapt-- this is particularly important, because unlike generic web crawlers, video crawling requires site-specific and semantic interpretation of a web-page's content.
- We need to deploy proprietary techniques of generic, statistical, template-based 'visual' crawling, specialized blog crawling techniques and also other proprietary tools that have been developed to incorporate site-specific understanding of web-content in a rapid and scalable fashion.
- In the context of social networking and sites supporting user-generated video, one important metric for the "value" of the crawling system is the breadth/depth of video clip coverage. Since any crawl system may never get to know the absolute state of the system/website it is crawling, (the actual number of videos present in the crawled site), a statistical estimate of coverage is needed to quantify the breadth/depth of coverage. This statistical coverage, (video clip coverage and user coverage), is measured on an ongoing basis, and the results are used to synthesize crawl schedules and maximize coverage.

In the context of term relevance and recallable information for internet video, it is

interesting to consider the impact of terms emanating out of clip attributes such as user comments, data from subtitles/closed caption of the video, and video-speech to text etc. Such information has the potential of introducing 'search term noise' -- misplaced emphasis on words and terms that are not likely to be used by the user to recall the specific video. On the other hand, there are some situations where these terms may distinctly add value, for example, detecting either 'adult' or 'pornographic' content. Similarly, when one can compare audio and video signatures, speech processing techniques also help to identify and mitigate the proliferation of copyright content. When one considers long-tail content, speech processing is a double-edged sword -- in case of long tail videos that lack accompanying meta-data, speech processing in the audio track of the clip can sometimes enhance the search/recall function by enhancing the meta-data. However, it can also add irrelevant meta-content corresponding to sets of spoken words that are insignificant to the clip. The semantic understanding of media for "popular' material is, in the best case, useful for "research" purposes, (a journalist trying to find a specific utterance somewhere in a five minute clip), and in the worst case introduce lot of search noise.

V360 – A 360 DEGREE VIDEO ANALYTICS SOLUTION

The explosive growth in online video consumption provides a unique opportunity for Service Providers, (whose access networks are the source of this traffic), to gain a new understanding of consumer video behavior that can yield strategic as well as tactical actionable intelligence. What Service Providers have is reams and reams of video (TV/VOD/Internet Video) related tune-streams/click-streams. Fundamental differences in internet browsing on the one hand, and consumption of video on linear TV, VOD or internet streaming on the other, require specialized analytic tools that go far beyond standard internet analytics such as sites visited, time spent, uniques etc. to obtain actionable intelligence regarding video consumption habits. An analytics product must enable its users to very easily, meaningfully and flexibly zoom in on arbitrary subsets and views of the database, (in this case, the database of video tune-events, click-streams, users, programs etc.), get meaningful statistics about that subset, mine correlations and relations between different subsets and do all of this in the most user-friendly, intuitive and scalable fashion. The more domain knowledge incorporated into the database and analytics model, the more meaningful are the tools and results presentation in the analytics product. The semantic knowledgebase contributed by the Universal SmartTag Technology, coupled with tremendous information from the clickstreams representing end users' content navigation behavior, constitute the basic pillars of a semantic video analytics product that provides true insights into consumer behavior.

Unique Positioning And Expertise For Video Analytics Product

The V360 platform's uniqueness is that its video-specific analytic tools are based on a deep understanding and proven expertise in two principal aspects of video consumption by users.

The first aspect consists of the specifics of the service provider *video consumption environment*: Veveo's analytics model incorporates the different consumption modes, user-habits and business relationships involved in the video delivery system. Video can be consumed by a user via basic linear TV, time shifted DVR, specialized channel packages of linear TV, PPV, free or paid VOD, as well as many sites on the internet

via PCs, Mobiles or internet set top boxes. Temporal patterns of video consumption based on time of day, week-days versus week-ends, as well as time-shifted and broadband video are the outcome of decades of conditioning to a dominant television culture and more recently convergence, time shifting mechanisms and internet video as a phenomenon. The influence of an interactive TV guide design on the accessibility of a program is significant. Video consumption is also significantly influenced by access networks and behavior exhibited by consumers related to the device platform -- short clips are preferred on mobile platforms, interactivity is the preferred embodiment for desktop computers, and passive consumption dominates TV. In the context of TV, VOD, and internet streaming, new business relationships between content creators, content distributors and service providers need to be, and are currently being, established.

The second unique aspect of V360 is its fine-grained characterization of what videos a user is watching. In this respect, the V360 analytics platform provides a continuously evolving *fine-grained semantic taxonomy* potentially at an individual level, of the hundreds of millions of video assets available in archives, on air, and available from on-demand servers, across internet sites. The semantic classification of TV/VOD programs, movies and internet videos into micro-categories and millions of tags and clusters is based on Universal Smart Tag (UST) technology and the V360 content database engine, and goes well beyond the limited TV-program segmentation that stems from classic TV analytics regimes. A fine-grained classification engine based on millions of Smart Tags becomes necessary especially for analyzing and understanding consumption in destination sites such as Youtube, DailyMotion and so on -- each of these can be deemed a 'video internet' in and of itself, and each of these sites' millions of uploaders is potentially a distinct channel with its own flavor. The Universal Smart Tag technology is a core component of Veveo's commercially deployed vTap internet product, which services millions of video streams and page hits to internet-enabled phones on a daily basis.

Applying Video Expertise In V360 Analytics

As mentioned previously, a business analytics engine must arm the user with the most meaningful domain dependent tools in order for the user to naturally filter and sift through raw data and mine actionable information out of that data. V360 tools based on the video consumption environment enable a focus on specific subsets of programs based upon: which media companies (with whom the service providers need to negotiate) produce the content, channels airing the content, channel schedule attributes such as airtime (time of day, day of week), LIVEness, VOD attributes such as paid or subscription, usage based filters (programs viewed by at least some number of viewers) and viewer subsets (programs viewed by consumers who spend on paid VOD content.)

V360 also provides ways to filter programs based not only on the program meta-content (i.e. genres, cast members and types of programs available in TV and VOD listings), but also on the Smart Tags associated with the programs. Enhancing the meta-content enhances the tools and filters for selecting programs using V360. In addition, V360 enables sifting through and understanding at a higher level the typical habits of users, for example, do people who buy subscription packages on TV typically watch new and paid movies on VOD? It also enables understanding clusters of users whose viewing habits are similar, (watching similar TV programs or movies), but in two distinct partitions – one cluster using VOD/Pay per

view/Premium Subscription and the other cluster using only free/basic subscription; identifying clusters of users who can be targeted in a marketing campaign that promotes certain kinds of subscription packages based on their affinity to their long tail interests, etc.

Arbitrary subsets (clusters) of users and programs can thus be easily specified in V360 and the relevant statistics for those subsets can be browsed. Additionally, V360 can correlate events in different subsets to uncover new relationships between entities.

The V360 platform provides tremendous flexibility in defining business rules. It also allows Service Provider personnel to pose specific pointed queries based on details of their own network's video consumption environment, and on rich fine-grained UST-based semantic classification, a capability that cannot be replicated in general purpose analytics tools.

Scalability And Trading Off Accuracy And Response Times By Sampling

The V360 analytics correlation engine is based on a distributed computing backend architecture that scales to peta-bytes of data, and then briskly sifts through vast tables of Service Provider video consumption data to deliver the desired analytics information. In addition to enabling video events to be visualized and comprehended at a higher level, the V360 platform also detects various patterns-- examples include Content Affinity Clusters, (related videos watched by users if they watched a specific topic), User Affinity Clusters, (a cluster of users with similar video behavior), Language/Ethnicity Affinity clusters etc.

One of the unique aspects of the V360 platform is its ability to support analysis of the data in a completely deterministic fashion, by considering the entire data set, or in a stochastic fashion by considering a subset of the original data set. The subset is sampled from the original data set using a sophisticated set of statistical sampling algorithms that are fine tuned to the idiosyncrasies of the underlying domain, while guaranteeing accuracy within a specified bound.

Although the analytics engine is highly scalable, sometimes speed or response time is more desirable than exact accuracy. V360 enables the user to choose among various progressive modes that trade accuracy for speed by using statistical sampling techniques that guarantee bounded errors. The statistical sampling techniques provide a way for the user to obtain a quick answer about a very large data set, (for example, two years worth of viewing habits for a very large metro), while a more accurate answer can be scheduled to be computed at a later time.

V360 Insight: Intuitive And Hands-On Analytics

Combining multi-touch development platforms such as the Microsoft Surface, with V360's domain specific database and analytics platform, Veveo is building an exciting product, **V360 Insight,** that enables executives and business owners to directly and intuitively interact with user behavior data on multiple dimensions and get responses to ad hoc arbitrary queries. V360 Insight empowers executives with real-time access and control over massive customer data, without requiring the IT department to create new reports each time a new query needs to be answered. Queries can be dynamically composed and analyzed through intuitive visualization, without any third party involvement. Veveo believes such an interactive and direct hands-on system will dramatically enhance an executive's understanding of customer viewing behavior

with respect to programs, channels/networks, media companies etc.

Platform Aspects

The V360 product is a platform that enables several diverse applications that can be built by third parties based on analytics data and information. The API exposes the semantic information at various levels of granularity across Space-Time for applications such as: Data Visualization, TV/VOD promotions, TV and VOD Recommendations, TV Program popularity and ratings based on exact counting, Targeted and Personalized advertisement insertion, and so on.

Conclusion

As connected multimedia devices proliferate in the next several years, the ability to easily and seamlessly discover and consume media from the TV and VOD world on the one hand, and the internet video world on the other, will be one of the primary factors that determine end user adoption of videos and video services. Veveo believes its vTap technology will enable service providers to meaningfully engage the 'on-demand' consumer generation as it taps the plethora of video across a multitude of input and display constrained devices.