2015第四屆海峽兩岸高校高端IT人才論壇

非結構化大數據智慧分析技術及應用

北京大學　信息工程學院

朱躍生 教授

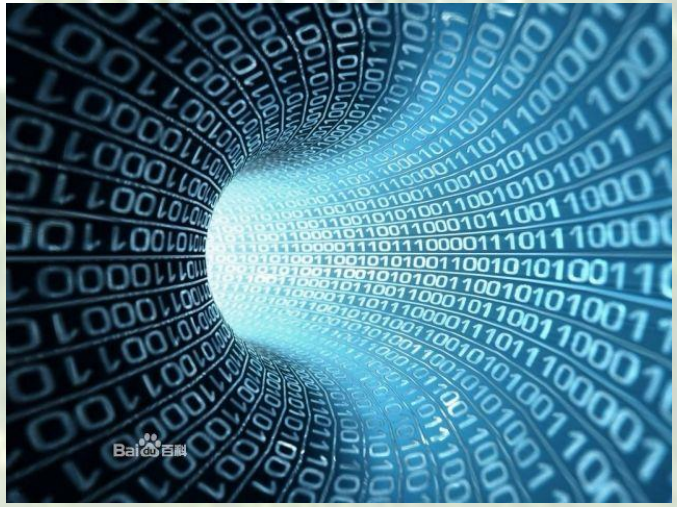
2015年7月27日

台中

北京大學

内容提要

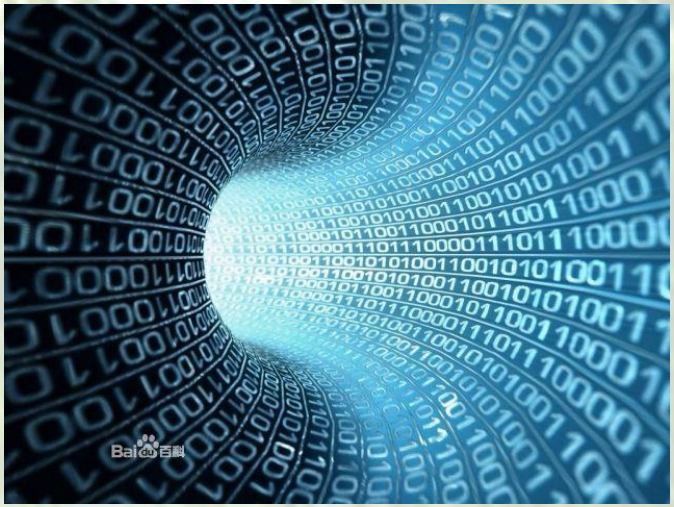
- 大數據基本概念
- 大數據理論基礎
- 非結構化數據智慧分析技術
- 大數據的安全問題



北京大學

内容提要

- **大數據基本概念**
- 大數據理論基礎
- 非結構化數據智慧分析技術
- 大數據的安全問題



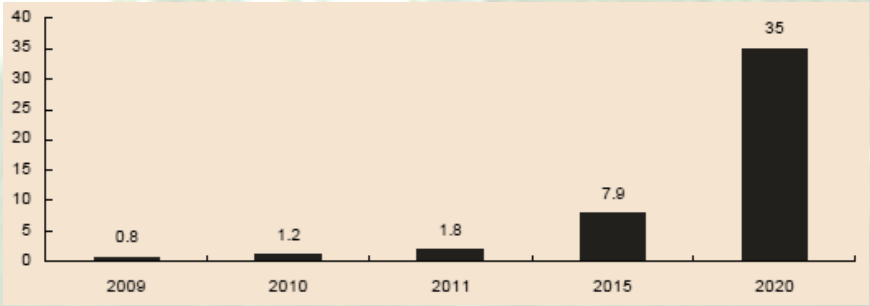
北京大学

大數據形成(big data)

◘ 每天產生巨大數據

- 互聯網（社交、搜索、電商、微博）、物聯網（感測器，智慧地球）、車聯網、GPS、醫學影像、安全監控、金融（銀行、股市、保險）、電信（通話、短信）正在瘋狂產生著數據

Email：全球發送約 3 百萬封 / 秒
Youtube：約3萬個小時視頻 / 天
推特：發佈 5千萬條消息 / 每天
亞馬遜：產生 6.3 百萬筆訂單 / 每天
Facebook： 7千億分鐘/月，
移動互聯網使用者發送和接收的數據高達1.3EB
Google ：處理24PB 數據 / 每天



IDC 預測，產生的數據量呈指數級增長，約翻一番/兩年，近兩年產生的數據量相當於之前產生的數據總量。

到2015年，全球數據量約達到7.9 ZB（Zetta-Bytes，1 ZB＝2^{70} byte， 10^{21} byte ）

✓ **已經遠遠超越了目前所能處理的能力！！！**

北京大學

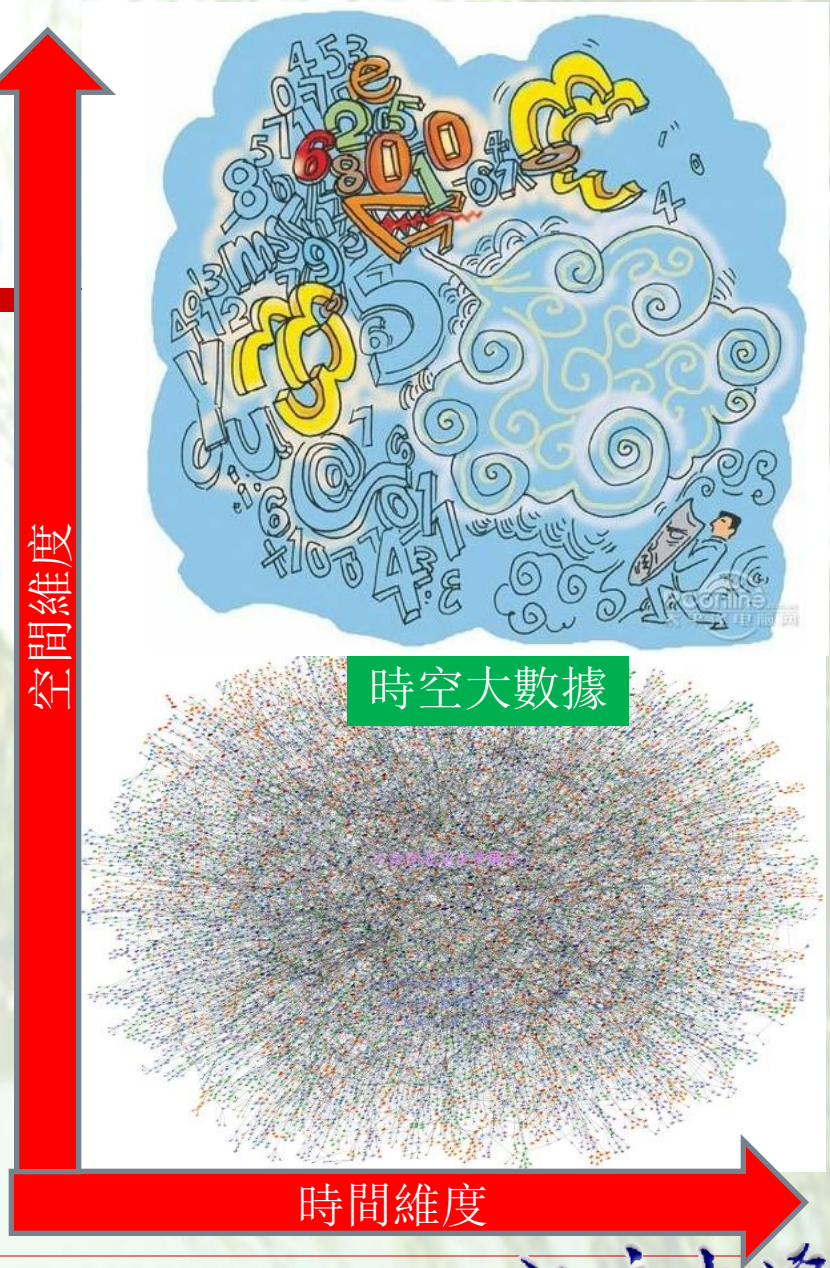
大數據(big data)

✓ 定義

數據量大到超出目前傳統數據庫軟體工具，在合理時間內達到獲取、管理、處理、並分析整理成可決策數據的能力

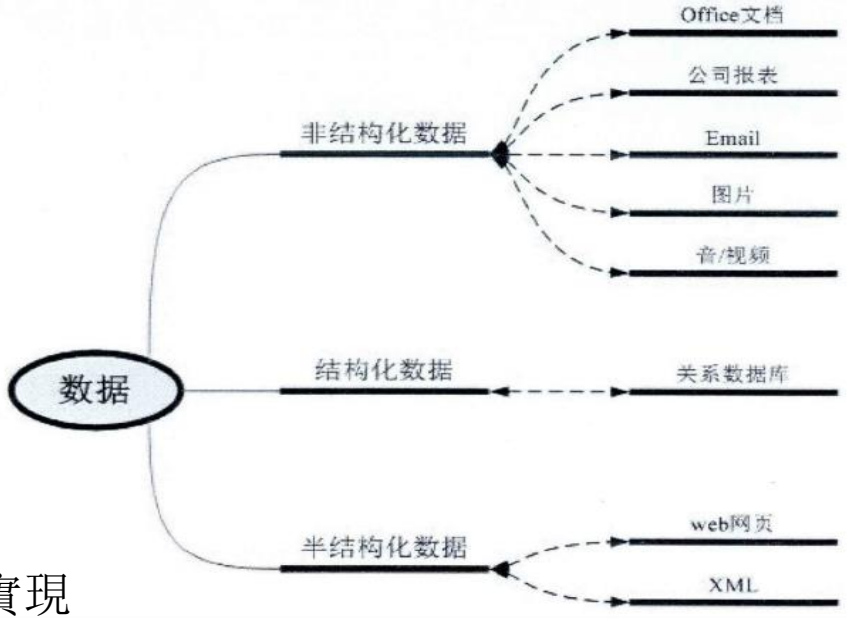
"Big data refers to data sets whose size is beyond the ability of typical database software tools to capture, store, manage and analyze."

- The McKinsey Global Institute, 2011



北京大學

大數據類型



- **結構化數據**

存在數據庫，可用二維表結構來邏輯表達實現

先有結構、再有數據

- **非/半結構化資料**

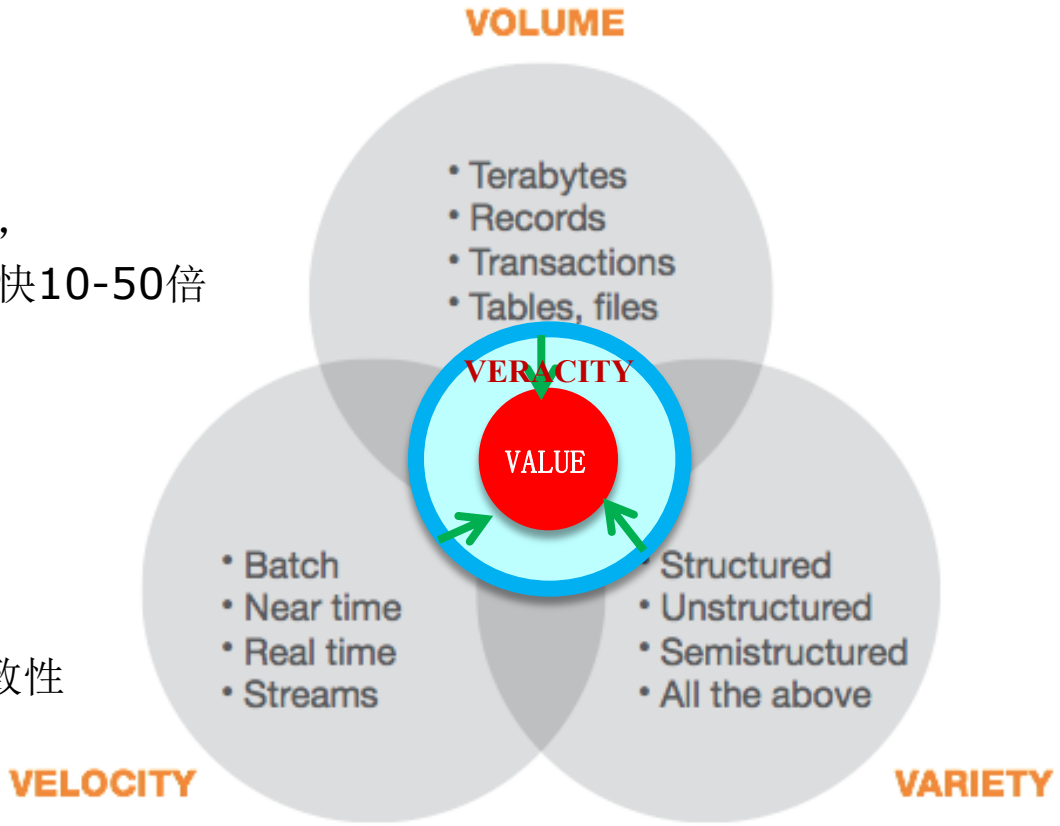
字段長度可變，辦公文檔、文本、圖片、XML、HTML、各類報表、圖像和音頻/視頻數據等

先有數據，再有結構

隨著網路技術的發展，非結構化數據的數量日趨增大

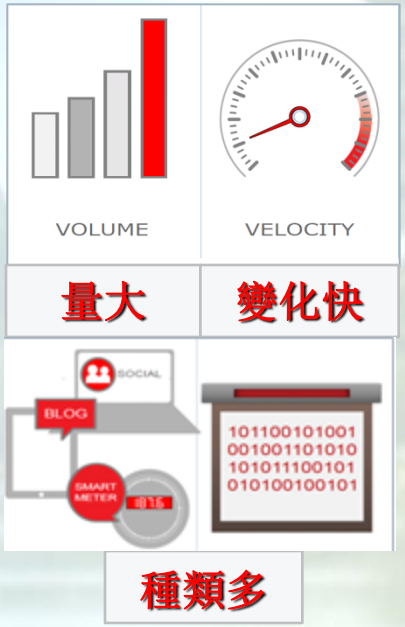
主要特點-5V

- **Volume（量大）**
  ZB級，非結構化數據大規模增長，
  占總量80%，比結構化數據增長快10-50倍
- **Velocity（變化快）**
  即時, 監控
- **Variety（種類多）**
  文本、圖像、音視頻、機器數據
- **Veracity （真實性）**
  完整性、模糊 / 隱性 》關聯一致性
- **Value（價值）**



挖掘 ⇨ 預測，諮詢，報告 ⇨ **效益**

基本屬性- 3I



Inexact
不確切

Inductive
可歸納

Incremental
增量

智能數據分析挖掘

價值

- 事物的基本特性
- 查找事物之間關聯性
- 離線大數據分析，即時動態處理
- 快速、精確預測出宏觀趨勢

大數據產生

大數據 ⇨ 海量數據 + 複雜類型數據+ **傳輸及處理方法**

- **海量交易資料**

交易數據和連線分析數據，結構化、靜態、歷史數據

- **海量交互數據**

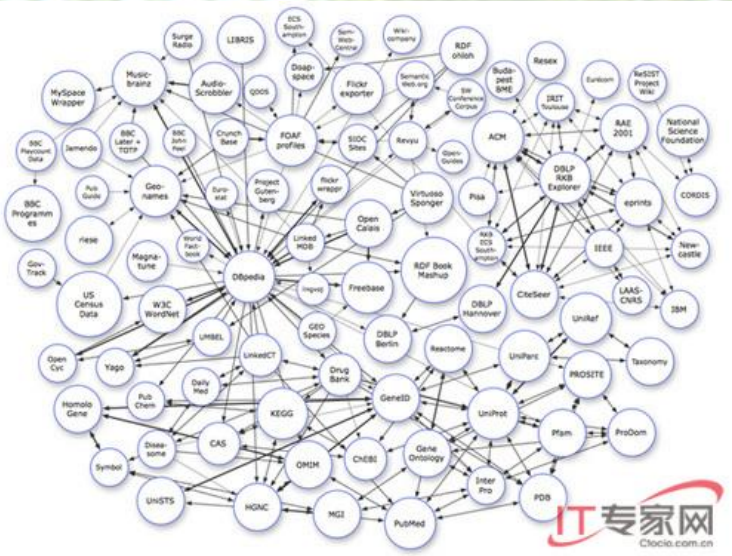
社交媒體數據（Facebook、Twitter、LinkedIn，微信，微博），呼叫記錄、設備和感測器數據、GPS和地理定位映射數據、海量影像檔、Web文本、電子郵件.....等等

- **海量數據傳輸及處理**

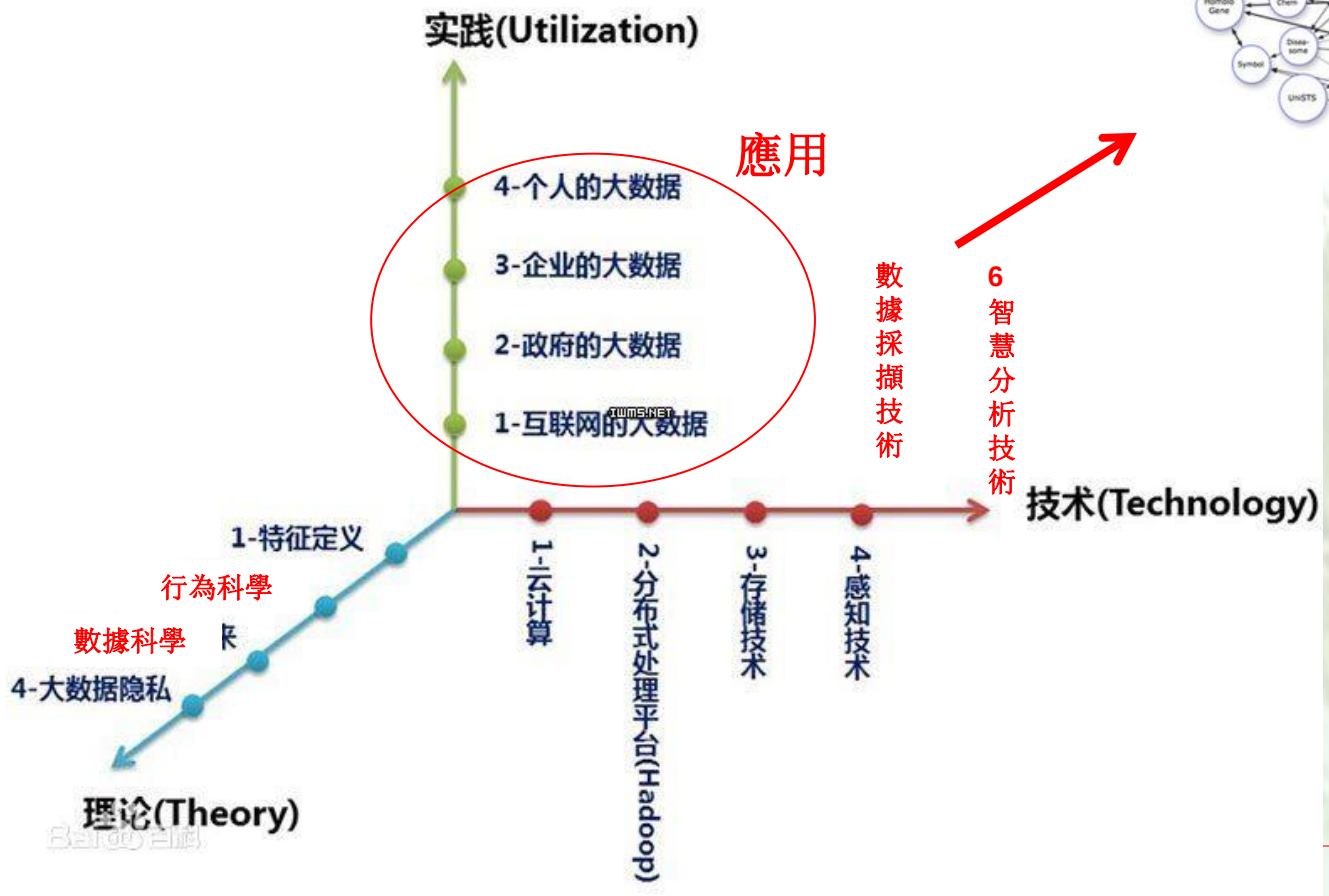
數據密集型處理架構，分析演算法，存儲方法，容災備份，修復方法，傳輸頻寬及方法，QoS/QoE

北京大學

大數據的理論與實踐

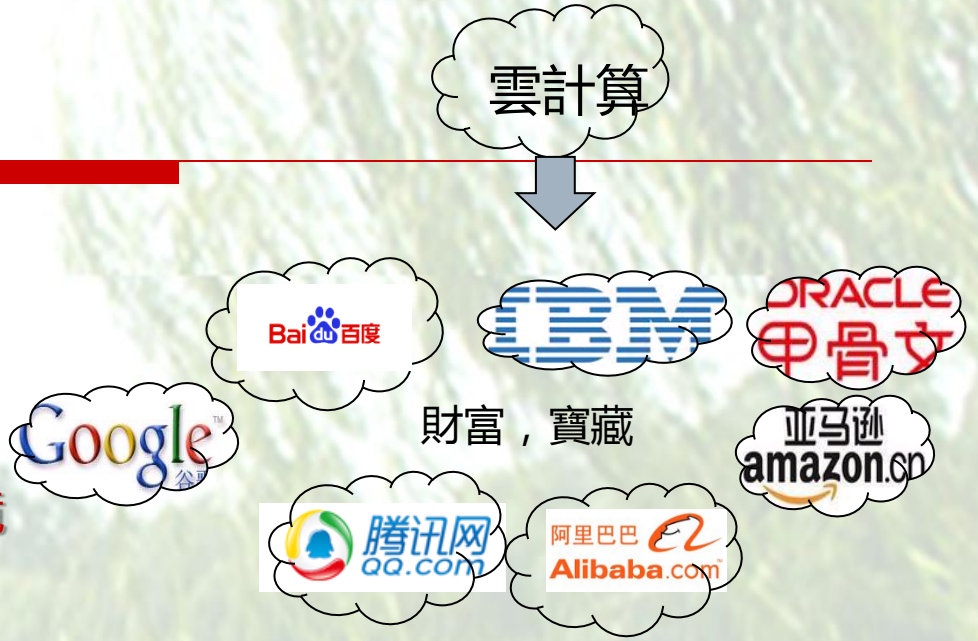


關聯性

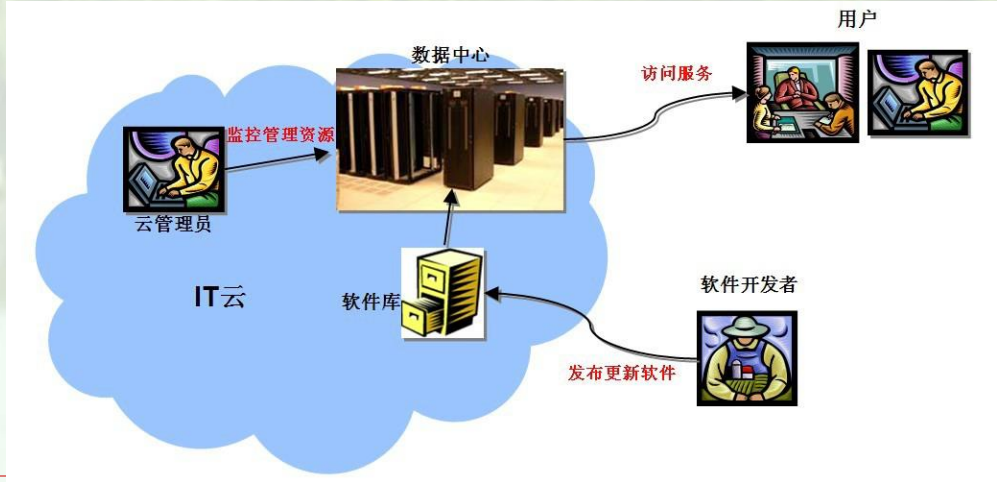


北京大學

大數據與雲計算



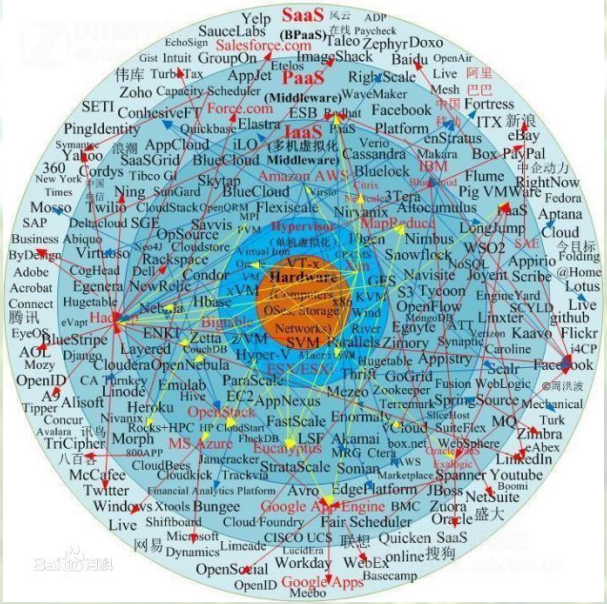
- ✓ 數據：資產源
  雲為資產提供存儲、訪問和計算環境
- ➢ 雲服務
  海量存儲和計算，如視頻智能監控
- ➢ 挖掘價值性信息和預測性分析
- ➢ 政府、企業、個人決策和服務



北京大学

技術架構挑戰

網路架構、數據中心、運維挑戰



➢ 現有數據庫管理技術

- ✓ 傳統數據庫不能處理數ZB/TB 級別的數據
- ✓ 不能支援高級別數據分析
- ✓ 急速膨脹的數據體量超越傳統數據庫的管理能力

➢ 經典數據庫沒考慮數據的多類別，沒考慮非結構化資料描述及管理

➢ 即時性挑戰

北京大學

大數據相關技術

➢ 分析技術

- 數據處理：自然語言處理
- 統計和分析：排行榜；文本分析
- 數據採擷：關聯規則分析；分類；聚類
- 模型預測：預測模型；機器學習；建模仿真

➢ 相關技術

- 數據獲取：工具
- 數據存取：數據庫
- 基礎架構支持：雲存儲；分布式系統
- 計算：雲計算

➢ 存儲

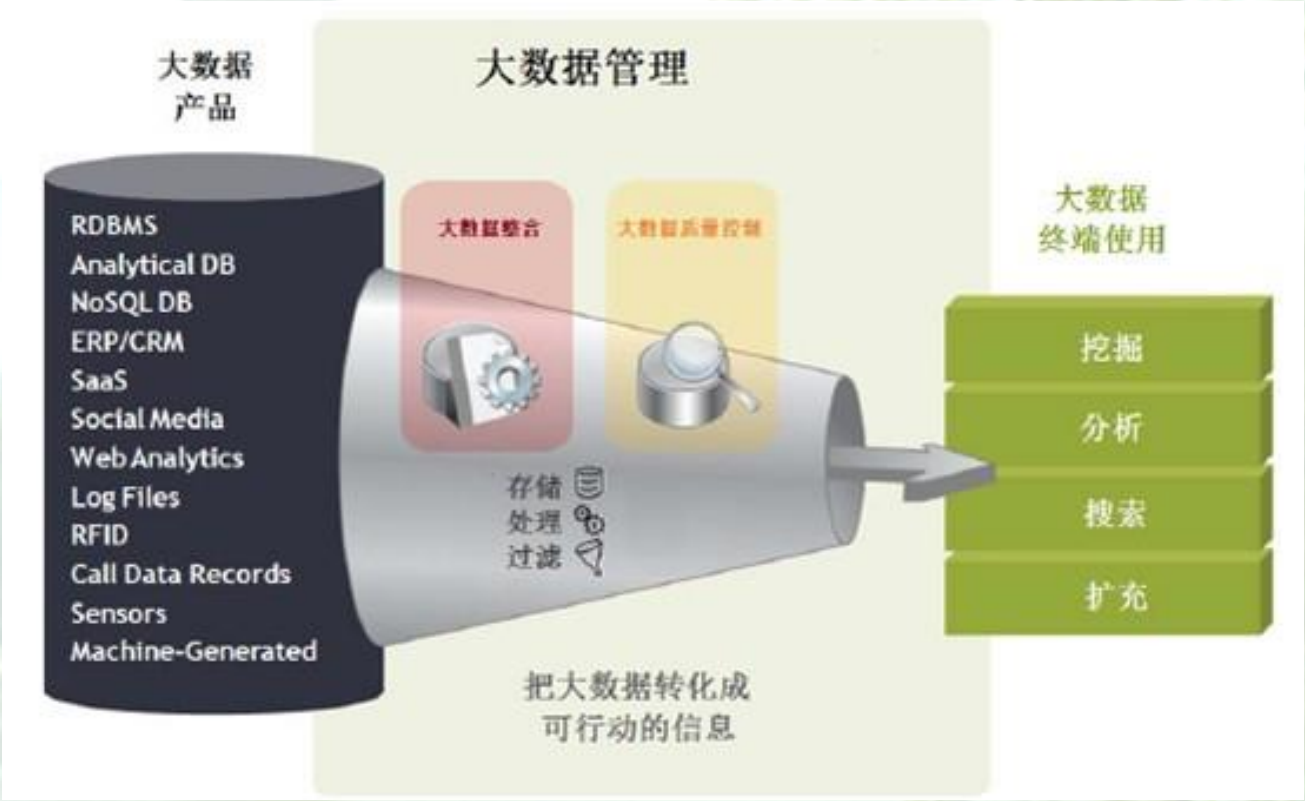
- 結構化數據：
  海量資料的查詢、統計、更新等操作效率低
- 非結構化數據
  圖片、視頻、word、pdf、ppt等檔存儲
  檢索、查詢和存儲
- 半結構化數據
  轉換為結構化存儲
  按照非結構化存儲

➢ 解決方案

- Hadoop（MapReduce技術）
- 流計算

北京大學

技術與創新架構設計

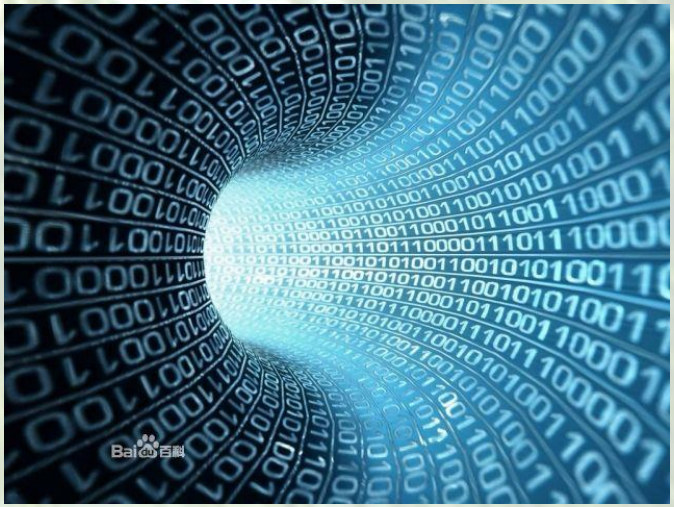


通過快速（velocity）的採集、發現和分析，從大量（volumes）、多類別（variety）的資料中提取較真實（Veracity ）的價值（value）

北京大學

内容提要

- 大數據基本概念
- **大數據理論基礎**
- 非結構化數據智慧分析技術
- 大數據的安全問題



北京大学

大數據的處理

- **現有及歷史數據的處理**

  存儲，格式轉化，分析，描述，挖掘，檢索，查詢，安全

- **正在採集的數據處理**

  傳輸，QoS，存儲，格式化，分析，描述，挖掘，歸檔，檢索，查詢，安全

- **未來產生數據的處理**

  預測，傳輸，QoS，存儲，格式化，分析，描述，挖掘，檢索，查詢，安全

北京大學

大數據處理的分類處理

- **結構化數據**

  在數據庫中，可用二維表結構來邏輯表達實現的數據

- **半結構化數據**

  不用數據庫二維邏輯表來表現的數據，如XML、HTML、各類報表等

- **非結構化數據**

  字段長度可變，如文本、圖像、音視頻等數據

北京大學

大數據處理的理論基礎

✓ **北京大學 數學學院 教授 鄂維南（E Weinan）院士 ：**

“數據科學所依賴的兩個因素是數據的廣泛性和多樣性， 以及數據研究的共性”

“數據科學主要包括兩個方面： 用數據的方法來研究科學和用科學的方法來研究數據”

- 用數據的方法來研究科學
  生物信息學、 天體信息學、数字地球等領域

- 用科學的方法來研究數據
  數據的獲取，存儲，和數據的分析

✓ **涉及領域**

数字信號處理、通信技術、計算數學、電腦科學、統計學、機器學習、數據採集、數據庫等

北京大學

數據類型及模型

（1） 表格。这是最为经典的数据。

（2） 点集（point cloud）。很多数据都可以看成是某种空间的一堆点。

（3） 时间序列。文本，通话，DNA序列等都可以看成是时间序列。它们也是一个变量（通常可以看成是时间）的函数。

（4） 图像。可以看成是两个变量的函数。

（5） 视频。时间和空间坐标的函数。

（6） 网页，报纸等。虽然网页或报纸上的每篇文章都可以看成是时间序列，但整个网页或报纸又具有空间结构。

（7） 网络数据。

还可以考虑更高层次的数据，如图像集，时间序列集，表格序列等等。

数据类型	模型
点集	概率分布
时间序列	随机过程（如隐式马氏过程等）
图像	随机场（如吉布斯随机场）
网络	图模型，贝叶斯模型

数學結構

特徵提取及分析

（1） 相关性。判断两组数据是不是相关的。

（2） 排序。比方说对网页作排序。

（3） 分类、聚类。把数据分成几类。

北京大學

數學結構處理

- **度量結構**
  在數據集上引入度量，及距離，形成度量空間

- **網路結構**
  引入某準則，建立連接，形成網路

- **代數結構**
  可把數據看成是向量，或矩陣，用代數方法表達

- **拓撲結構**
  不同的尺度去看數據集，得到的拓撲結構不一樣

- **函數結構**
  線性函數，用於線性回歸；分片常數，用於聚類或分類；分片多項式，如樣條函數；小波展開函數等

北京大學

數據分析面臨的挑戰

- **數據量大**
- **維數高**

  模型的複雜度和計算量，
  隨著維數的增加而指數增長

  - ✓ **限制數學模型**
    如假設概率密度遵循正態分佈
  - ✓ **利用數據的特殊結構降維**
    如稀疏性，低維或低秩，光滑性等

- **類型複雜**
- **噪音大**

大數據处理算法的需求

- **降低处理算法的複雜度**

  優化方法降低計算量

- **平行計算**

- **雲計算**

- **噪音大**

（1）k--平均（k-means）法
對數據作聚類的最簡單有效

（2）支持向量機
基於變分（或優化）模型的分類計算算法，側重整體趨勢

（3）期望最大化（EM）演算法
如基於極大似然方法（ maximum likelihood）的參數估計

（4）網頁排序計算算法PageRank
網頁排序由在互聯網中的重要性決定，把排序轉換成矩陣的特徵值

（5）貝葉斯方法
從先驗的概率密度，結合已知的數據得到後驗的概率密度

（6）k--最近鄰域法
用鄰域數據來作分類，側重局部資訊

（7）AdaBoost
通過變換權重，把弱分類器變成強分類器

（8）深度學習法

數據科學的知識教育體系

- **數學基礎知識**
  三大基礎：微積分、線性代數和概率論
  隨機過程、函數逼近論、圖論、拓撲學、幾何、變分法、群論等
- **計算機科學基本知識**
  計算機語言、數據庫、數據結構、可視化技術等
- **算法基本知識**
  數值代數、函數逼近、優化方法、網路算法、計算幾何等
- **數據模型**
  回歸、分類、聚類、參數估計等
- **專業課程**
  信號理論、數字信號處理、時間序列分析、影像處理、視頻處理、編碼理論、自然語言處理、文本處理、語言識別、通信系統、信息安全、網路技術、移動計算、雲計算等
- **其它應用專業課**
  生物信息學、地理信息學、金融數據，計算廣告學、推薦系統等

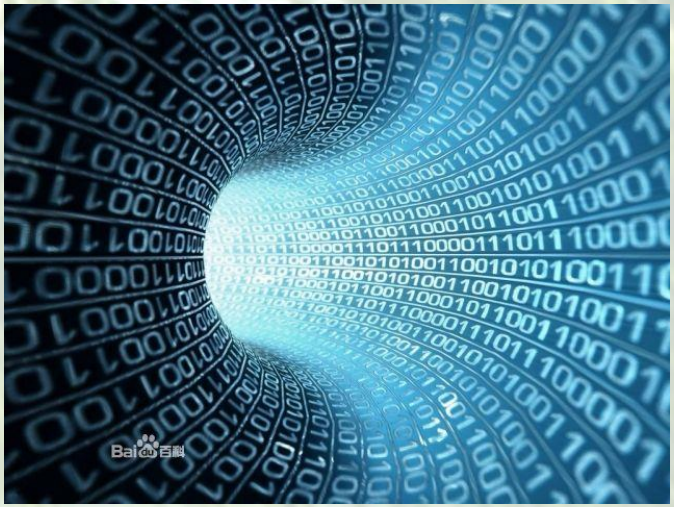
十大技能

1，大數據基礎知識
2，統計學
3，程式設計
4，機器學習
5，文本挖掘/自然語言處理
6，視覺化
7，大數據和雲計算
8，數據獲取
9，數據再處理
10，工具箱

北京大學

内容提要

- 大數據基本概念
- 大數據理論基礎
- **非結構化大數據智慧分析技術**
- 大數據的安全問題



北京大學

大數據處理架構


数据源
关键技术
结构化数据
非结构化数据
半结构化数据
数据清洗
语义分析和数据挖掘
数据集成
查询与索引
人机交互
数据存储
MapReduce及其它并发处理架构
可视化
数据传输、虚拟集群等其它支撑技术
数据采集与预处理
数据分析
数据解释

非結構化大數據管理技術

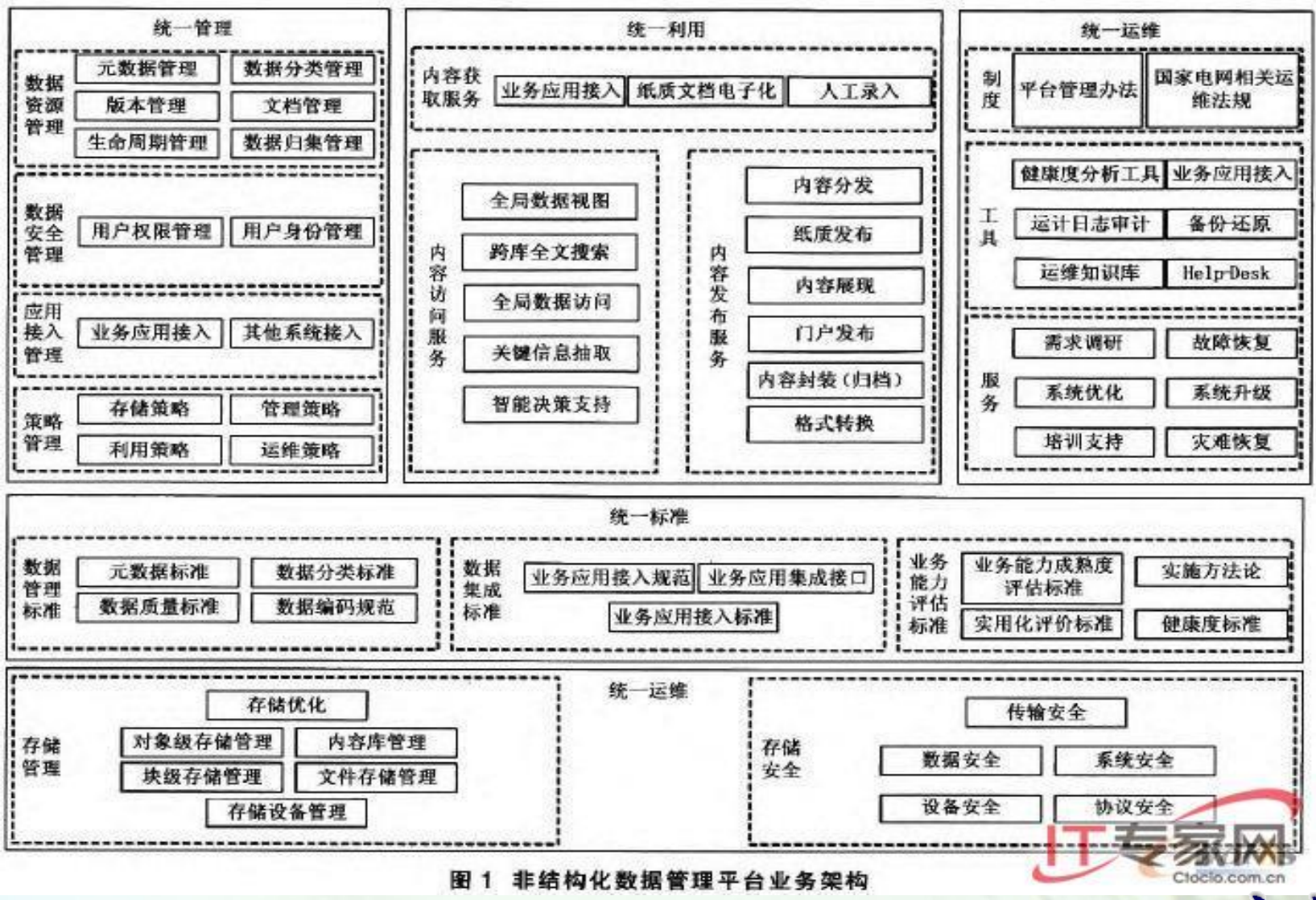
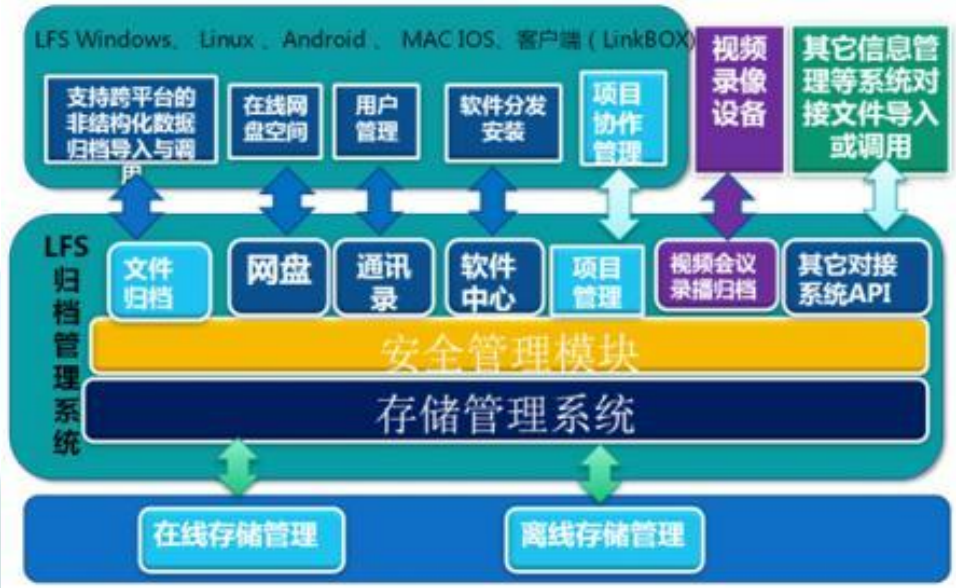
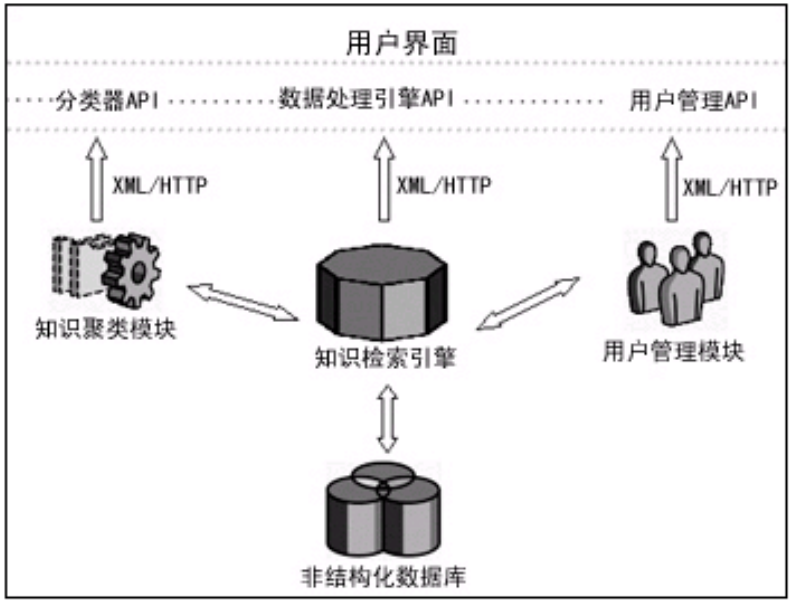


图 1 非结构化数据管理平台业务架构

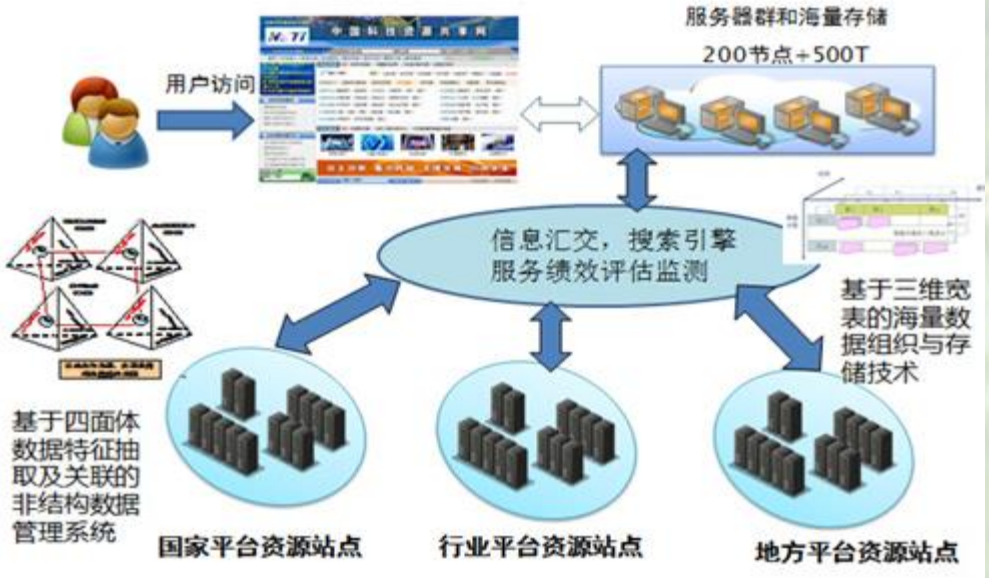
北京大学

非結構化大數據處理技術



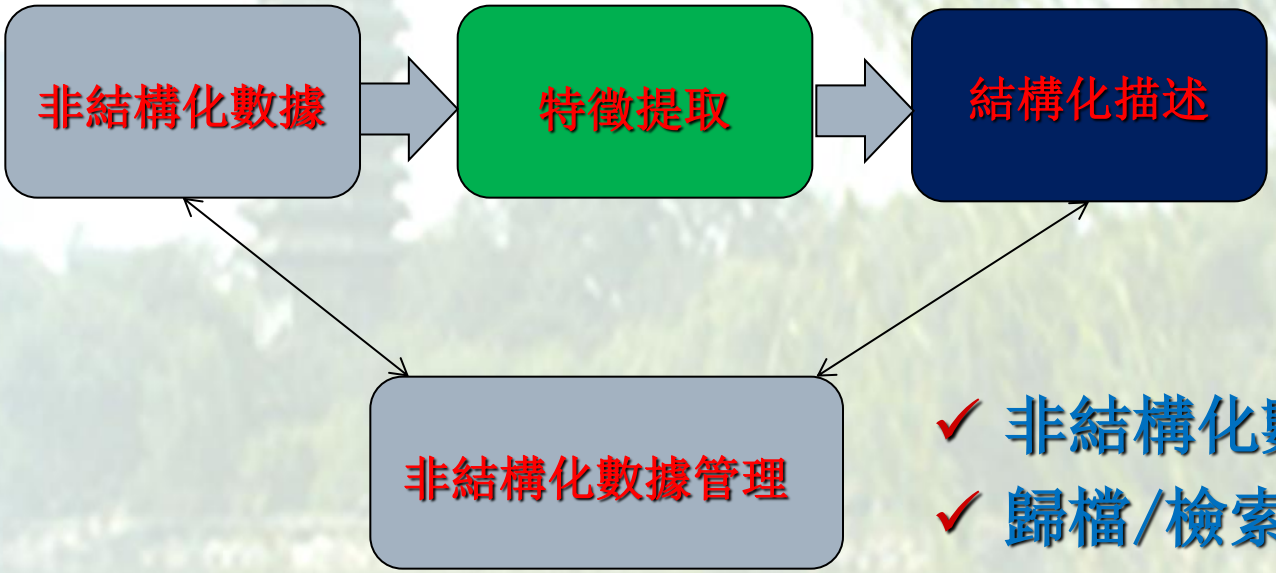


- ✓ 非結構化數據處理引擎
- ✓ 歸檔/檢索引擎
- ✓ 描述與分類與聚類



非結構化大數據的結構化描述

化解為結構化數據

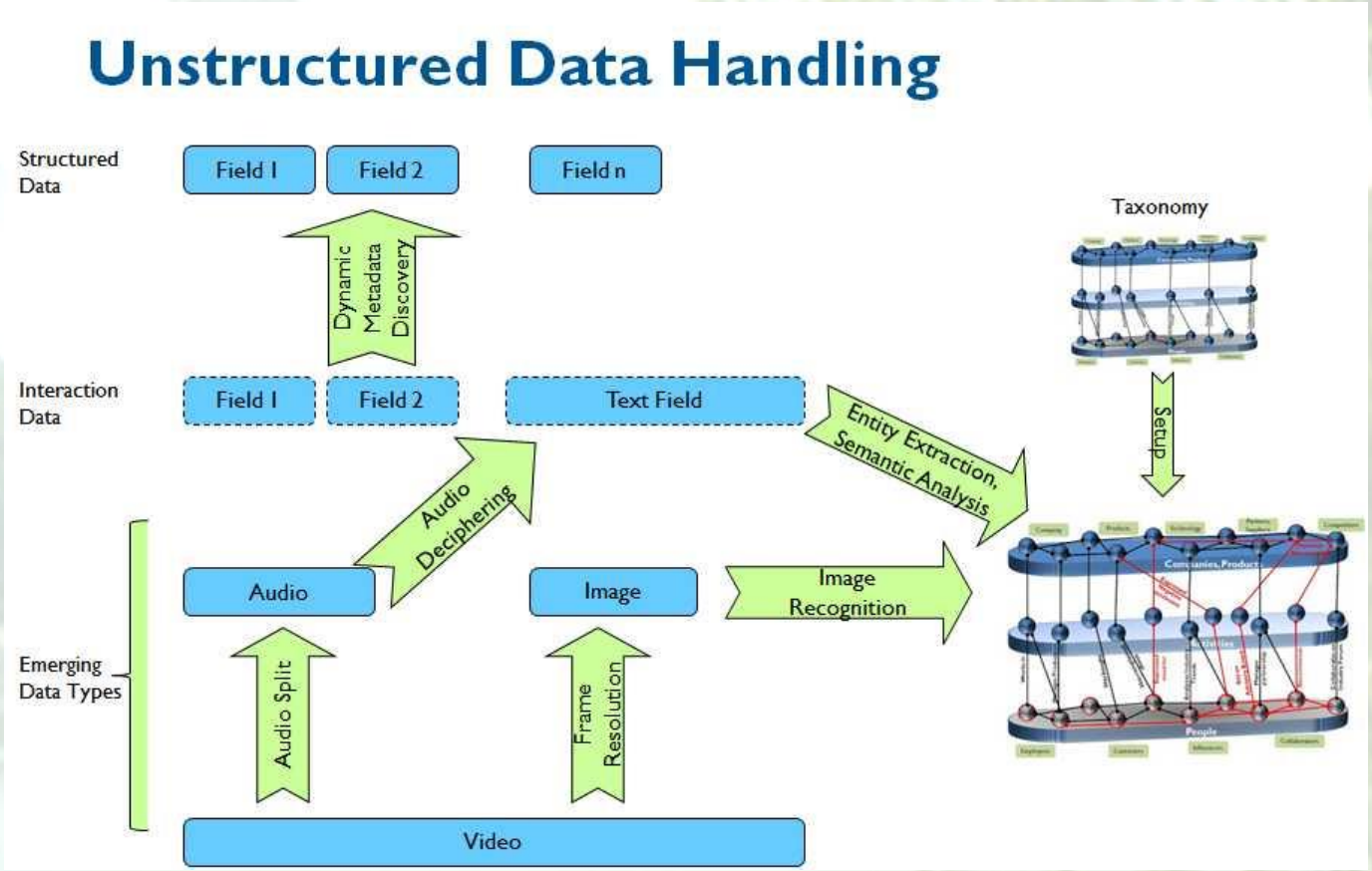


- ✓ 非結構化數據處理引擎
- ✓ 歸檔/檢索引擎
- ✓ 描述與分類與聚類

北京大學

非結構化大數據智能分析技術

視頻/圖像智能分析



音視頻/圖像智能分析

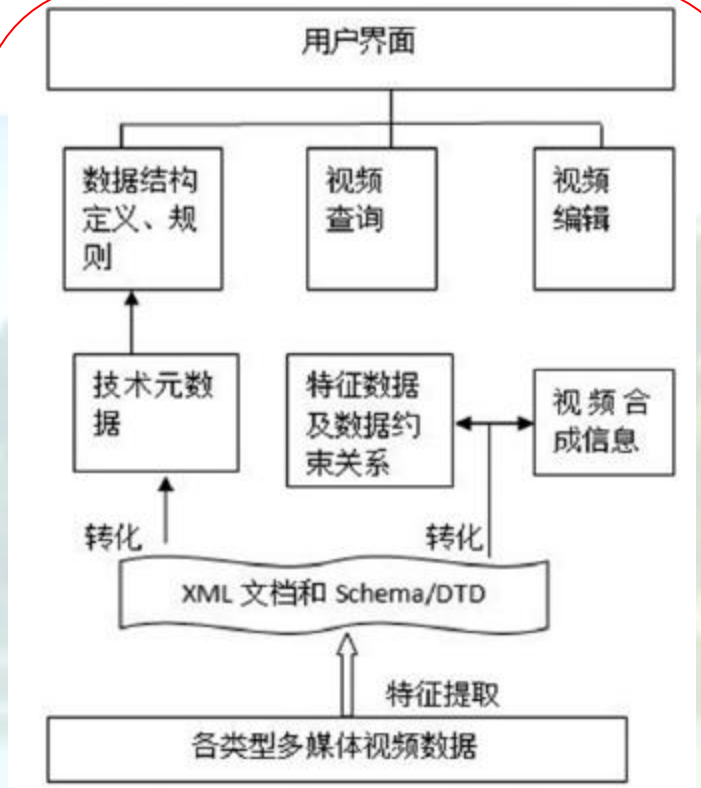
- ✓ 本質

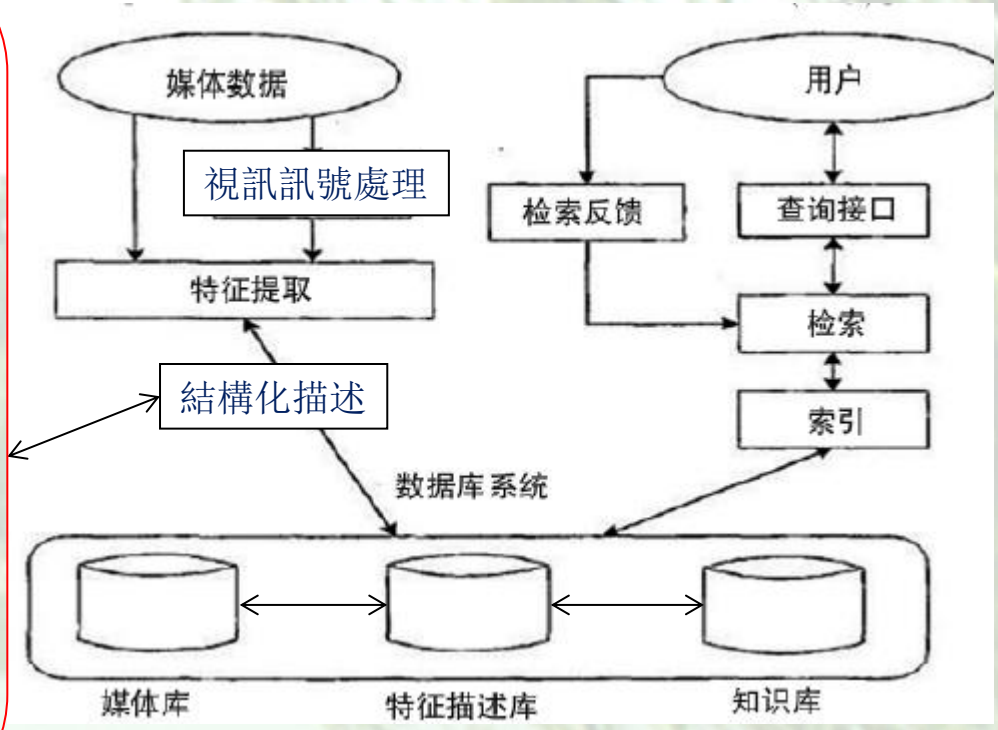
  動態信號分析，特徵提取，表達，描述，分類，檢索，管理，利用

- ➢ 音視頻/圖像指紋的提取
- ➢ 視頻智慧監控
- ➢ 智能動態檢測/跟蹤
- ➢ 離線/線上協同工作

北京大學

基於多媒體管理系統


用户界面
数据结构定义、规则
视频查询
视频编辑
技术元数据
特征数据及数据约束关系
视频合成信息
转化
转化
XML 文档和 Schema/DTD
特征提取
各类型多媒体视频数据


媒体数据
用户
視訊訊號處理
检索反馈
查询接口
特征提取
检索
結構化描述
索引
数据库系统
媒体库
特征描述库
知识库

音視頻/圖像指紋-數據鑒別技術

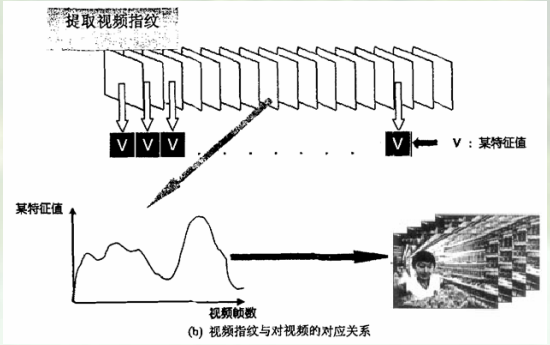
指紋技術 （Fingerprint）

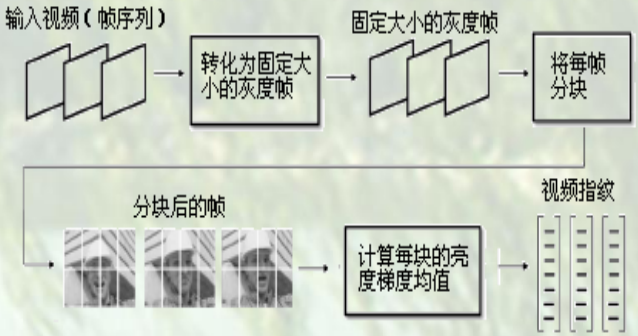
用指紋來鑒別人的身份

引伸 ⇒ **數字指紋**

數字視頻指紋

提取視頻內容的自身特徵，產生數字指紋，作為識別視頻信息內容唯一性判據

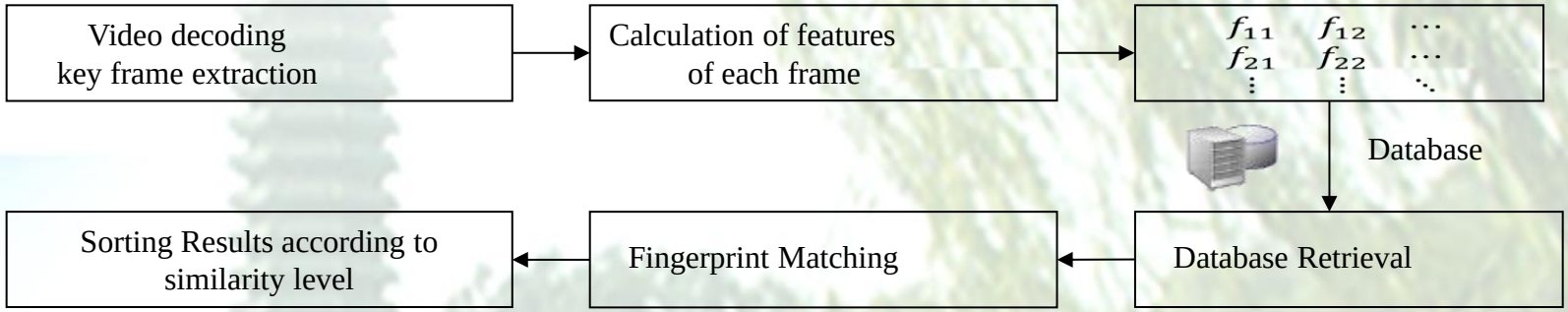




應用於視頻檢索和版權保護領域

北京大学

基於數字指紋技術平臺視頻管理系統



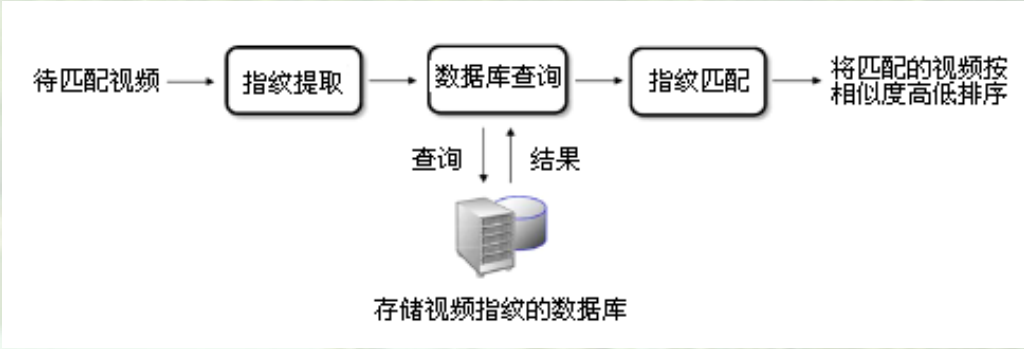

二等奖

TI DSP DESIGN CONTEST
OMAP INVITATION CONTEST

2009-2010
OMAP3组

参赛团队：北京大学
项目名称：一种基于数字指纹算法的视频管理系统
指导教师：朱跃生
团队成员：韩普晓、刘振羌、陈红阳、霍薇靖、余婷婷

Texas Instruments Semiconductor Technologies (Shanghai) Co.,Ltd.
德州仪器半导体技术(上海)有限公司　二零一零年五月



北京大学

數字媒體指紋/水印算法及管理系統

KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS VOL. 7, NO. 11, Nov. 2013 2754

A Robust Video Fingerprinting Algorithm Based on Centroid of Spatio-temporal Gradient Orientations

Ziqiang Sun[1], Yuesheng Zhu[1], Xiyao Liu[1] and Liming Zhang[2]
[1] Communication & Information Security Lab, Shenzhen Graduate School, Peking University
Shenzhen, Guangdong 518055 - China
[e-mail: sunzq87@gmail.com; zhuys@pkusz.edu.cn; lxyzoewx@163.com]
[2] Faculty of Science and Technology, University of Macau
Macao, 999078 - China
[e-mail: lmzhang@umac.mo]
*Corresponding author: Yuesheng Zhu

Received July 9, 2013; revised October 3, 2013; accepted Novemebr 10, 2013; published November 29, 2013

Abstract

Video fingerprints generated from global features are usually vulnerable against general geometric transformations. In this paper, a novel video fingerprinting algorithm is proposed, in which a new spatio-temporal gradient is designed to represent the spatial and temporal information for each frame, and a new partition scheme, based on concentric circle and rings, is developed to resist the attacks efficiently. The centroids of spatio-temporal gradient orientations (CSTGO) within the circle and rings are then calculated to generate a robust fingerprint. Our experiments with different attacks have demonstrated that the proposed approach outperforms the state-of-the-art methods in terms of robustness and discrimination.

Keywords: Content-based copy detection, video fingerprint, copyright control, spatio-temporal gradient, concentric rings

Multimed Tools Appl
DOI 10.1007/s11042-014-2073-4

A novel robust video fingerprinting-watermarking hybrid scheme based on visual secret sharing

Xiyao Liu · Yuesheng Zhu · Ziqiang Sun · Mengge Diao · Liming Zhang

Received: 30 November 2013 /Revised: 17 April 2014 /Accepted: 1 May 2014

Abstract In this paper, a novel robust Video Fingerprinting-Watermarking Hybrid Scheme (VFWHS) is proposed by applying fingerprinting technique to zero-watermarking algorithm for enhancing the performance of video Digital Rights Management (DRM). In the VFWHS, the content-based feature of the original video is extracted by an improved 3D-DCT fingerprinting algorithm and stored in a video fingerprint database for retrieval, while the watermark-based feature of the watermark image is extracted by a 2D-DCT fingerprinting algorithm for copyright identification. According to the visual secret sharing scheme, a master share is generated from the content-based feature while an ownership share is generated from the master share and the watermark-based feature. The ownership share is then stored in an ownership database. When a video is queried, its content-based features are generated, and a process of similarity measurement is executed to obtain the matched feature with the relevant ownership share. The copyright ownership can be identified by stacking the master share generated from the queried video with the relevant ownership share. Our experimental results demonstrate that the VFWHS can obtain the relevant ownership share precisely for reliable copyright identification in the mass content processing which cannot be achieved by the zero-watermarking algorithms, and provide explicit copyright identification to avoid possible copyright dissension which cannot be achieved by the fingerprinting algorithms.

Keywords Digital rights management · Visual secret sharing · Fingerprinting · Zero-watermarking · 3D-DCT

An Improved Fingerprint Algorithm of 3D-DCT for Video Fingerprinting

Mengge Diao[1], Yuesheng Zhu[1], Ziqiang Sun[1], Xiyao Liu[1], Limin Zhang[2]
1. Communication and Information Security Lab
Shenzhen Graduate School
Peking University, Shenzhen, China
2. Faculty of Science and Technology, University of Macau, Macao, China
Corresponding author's e-mail: zhuys@pkusz.edu.cn

***Abstract*—In this paper, a new learned basis set algorithm (3D-LBT) based on 3D-DCT (Discrete Cosine Transform) is proposed for video fingerprinting and matching, in which for different video categories an Adaboost-based machine learning method is applied to each category of videos for selecting suitable sets of 3D-DCT coefficients to generate fingerprints, and a weighted distance of fingerprints is also defined for fingerprint matching. Our experimental results have illustrated that the proposed algorithm outperforms the conventional 3D-DCT algorithm and the 3D-RBT (Randomized Basis seT) algorithm in terms of robustness and uniqueness. Moreover, the proposed algorithm has better security performance for copyright applications.**

***Keywords*—video fingerprint; copyright protection; machine learning; 3D-DCT; weighted distance; adaboost**

I. INTRODUCTION

With the development of internet technology video uploading and sharing gains increasing popularity through the network. Confronting the huge data of video flows, the problems such as video management, indexing, and copyright protection, begin to receive major concerns. Video fingerprint as a type of perceptive video hashing can efficiently solve these problems. Video fingerprint system consists of three parts: (1) Fingerprint extraction (2) Fingerprint searching (3) Fingerprint matching. In fingerprint extraction, video fingerprints are extracted from the query video using specified algorithms. In fingerprint searching, an index structure is used to find candidate fingerprints for matching with the database. Fingerprint matching proceeds through calculating the distance between fingerprints of the query video and candidate video in the database. And a threshold is given for judging whether the two videos are same to get the index results. The proposed algorithm in this paper makes improvements in the section of fingerprint extraction and matching, which will be specified in the following sections. Fingerprint searching part will not be focused in this paper.

Video fingerprints are usually applied in video identification [5], [6], [7], detection [2]-[4] and copyright protection [1], [5]-[7], thus fingerprints extracted from videos requires three properties: robustness, uniqueness and security. Robustness means that even video undergoes several transformations such as compression, brightness or contrast changes, fingerprints calculated from the transformed and original videos are similar enough to judge them as same. Uniqueness requires that fingerprints calculated from different videos should differ from each other to distinguish them even there are several similarities in video contents. Furthermore, security of fingerprints plays an important role in the application of copyright protection. Fingerprint security [1], [4]-[6] prevents videos from forgery attacks committed by muggers who attain legal rights for illegal videos through forging fingerprints.

Video fingerprints are content-based signatures derived from videos, thus features which are used to extract fingerprints are crucial. Considering the fusion of both temporal and spatial characteristics, spatio-temporal fingerprints [1], [6]-[9] can withstand not only geometrical attacks but also temporal attacks [1], [8]-[10], which leads to increasing attention to and pervasive applications of spatio-temporal fingerprints. Among these fingerprints, transformation based algorithms are a significant branch. One of the representative algorithms is the 3D-DCT (Discrete Cosine Transform) algorithm introduced by Baris Coskun in [1]. The algorithm selects the lowest 64 (4*4*4) 3D-DCT coefficients of videos as the features to extract fingerprints based on the theory that signal energy mainly exists in low frequency part. In order to improve the security of 3D-DCT based fingerprints, 3D-RBT (Randomized Basis seT) algorithm is proposed in [1]. In 3D-RBT, a random parameter is introduced to randomly select 3D-DCT coefficients. The parameter plays as a key to encrypt the selected DCT coefficients so that muggers cannot conduct forgery attacks without deciphering the key.

There are two problems existing in both 3D-DCT and 3D-RBT algorithm. (1) In 3D-DCT, the selected positions of 3D-DCT coefficients, which are also called bases, are fixed to the lowest 64 ones. Although the energy of video signals mainly concentrates in low frequencies, low frequencies are not just limited to the lowest 64 ones, those low frequencies outside the lowest 64 ones may also contain remarkable information. However, in 3D-RBT, random selections of 3D-DCT bases take the risk that the entire selected results may drop in the range of high frequencies, which will make the performance of the system degrade a lot. (2) As 3D-DCT transformation is open and the positions of the selected 64 coefficients are fixed, muggers can easily conduct two attacks: copyright steal through video fingerprints fabrication for illegal content; slight adjustments on coefficient selections which produce huge distinction of generated fingerprints from origin in order to skip

A Robust Audio Fingerprinting Algorithm in MP3 Compressed Domain

Ruili Zhou, Yuesheng Zhu

***Abstract*—In this paper, a new robust audio fingerprinting algorithm in MP3 compressed domain is proposed with high robustness to time scale modification (TSM). Instead of simply employing short-term information of the MP3 stream, the new algorithm extracts the long-term features in MP3 compressed domain by using the modulation frequency analysis. Our experiment has demonstrated that the proposed method can achieve a bit rate of above 95% in audio retrieval and resist the attack of 20% TSM. It has lower bit error rate (BER) performance compared to the other algorithms. The proposed algorithm can also be used in other compressed domains, such as AAC.**

***Keywords*—Audio Fingerprinting, MP3, Modulation Frequency, TSM**

I. INTRODUCTION

WITH the development of multimedia technology and the advent of massive music information, some applications like music identification, music copyright verification, music searching and audio monitoring are developing as well. Audio fingerprinting is a compact signature based on the content of audio signal, which represents an important acoustic feature and the essence of the music signal. Audio fingerprinting on raw audio format has been studied [1]. A block diagram [2] of audio fingerprinting in uncompressed domain is shown in Fig.1, which extracts audio fingerprint from the energy difference between two adjacent bands and can resist most of the distortions except time scale modification (TSM) above 2%. A audio fingerprinting in entropy domain is proposed in [3], its robustness to low pass filtering and @32kbps compression is better than that of [2], but other distortions are not considered in this method. A noise robust fingerprinting method [4] uses long-window analysis strategy to resist some unknown distortions. The time-frequency variations based on a transformation with efficient time-scale localization is analyzed in [5] and two fingerprints are created for authentication and recognition purposes, respectively. In [6], time-frequency theory integrated with psychoacoustic results on modulation frequency perception is used for audio fingerprinting. It not only contains short-term information about the signals, but also provides long-term information representing patterns of time variation. A cross entropy approach is considered to classify music signals and a better performance under TSM distortion is achieved.

As compressed audio signals are increasing and have become the dominant fashion of audio file storage in personal electronic equipments and transmission on the Internet, it is important to extract the audio feature in compressed domain directly. It is noted that only a few work of music retrieval in compressed domain are reported and most of the algorithms are directly transplanted from those in raw data domain. Modified Discrete Cosine Transform (MDCT) coefficients and their energy derivation are frequently used to extract the fingerprint. Fig.2 shows the general procedures of audio fingerprinting in compressed domain. A robust compressed domain audio fingerprinting algorithm is developed in [7], which takes the ratio between the sub-band energy and the full-band energy of a segment as intra-segment feature and difference between continuous intra-segment features as inter-segment feature, and robustness to some distortions except TSM is shown. Compressed-domain spectral entropy is utilized as the audio feature in an audio fingerprinting algorithm [8], but it can resist TSM up to 2% only. In [9]- [11] the beat, rhythm and tempo information of the music signal is extracted respectively. A video retrieval system(VRS) is developed in [12] for Interactive-Television in AC3 domain, in which long-term logarithmic modified DCT modulation coefficients(LMDCT-MC) are used for audio indexing and retrieval and a two-stage search (TSS) algorithm for fast searching is also proposed.

Fig.1 Block diagram of audio fingerprinting in uncompressed domain

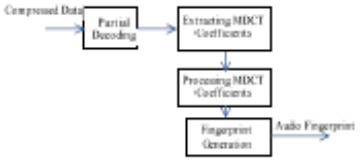


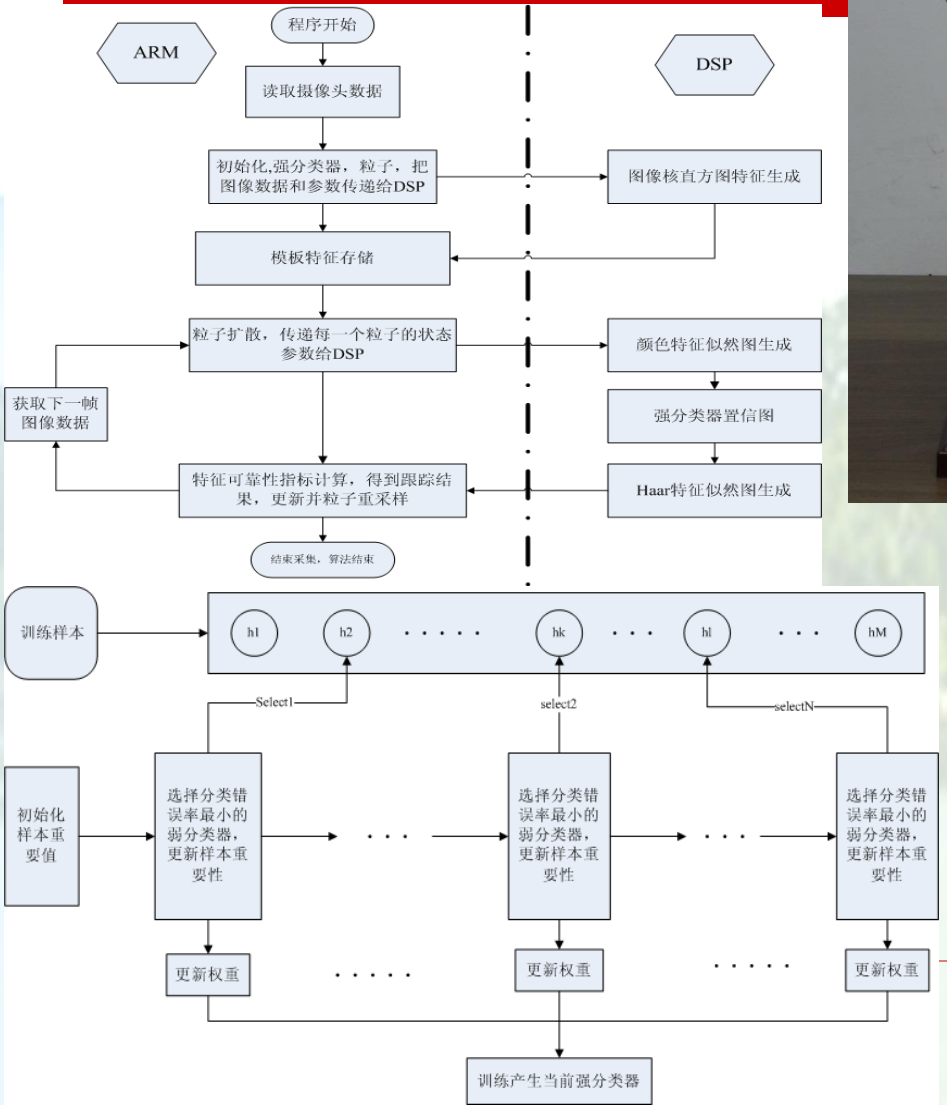
Fig.2 Block diagram of audio fingerprinting in compressed domain

In summary, most of the current algorithms in compressed domain are fragile to TSM distortions. It is our motivation to design an algorithm in compressed domain robust to TSM. Therefore, a new robust audio fingerprinting algorithm in MP3 compressed domain is proposed with high robustness to time

北京大学

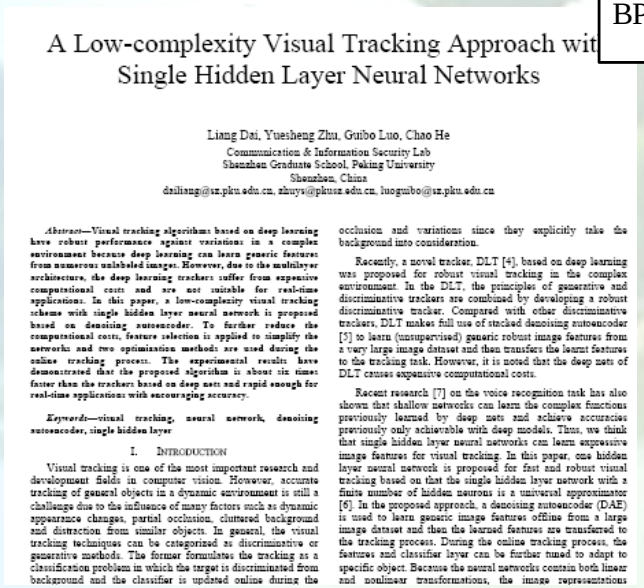
智能視頻分析

基於多線索融合的目標跟蹤算法



離線/線上協同工作

- 離線深度學習
- 線上精准匹配/跟蹤



A Low-complexity Visual Tracking Approach wit
Single Hidden Layer Neural Networks

Liang Dai, Yuesheng Zhu, Guibo Luo, Chao He
Communication & Information Security Lab
Shenzhen Graduate School, Peking University
Shenzhen, China
dailiang@sz.pku.edu.cn, zhuys@pkusz.edu.cn, luoguibo@sz.pku.edu.cn

Abstract—Visual tracking algorithms based on deep learning have robust performance against variations in a complex environment because deep learning can learn generic features from numerous unlabeled images. However, due to the multilayer architecture, the deep learning trackers suffer from expensive computational costs and are not suitable for real-time applications. In this paper, a low-complexity visual tracking scheme with single hidden layer neural network is proposed based on denoising autoencoder. To further reduce the computational costs, feature selection is applied to simplify the networks and two optimization methods are used during the online tracking process. The experimental results have demonstrated that the proposed algorithm is about six times faster than the trackers based on deep nets and rapid enough for real-time applications with encouraging accuracy.

Keywords—visual tracking, neural network, denoising autoencoder, single hidden layer

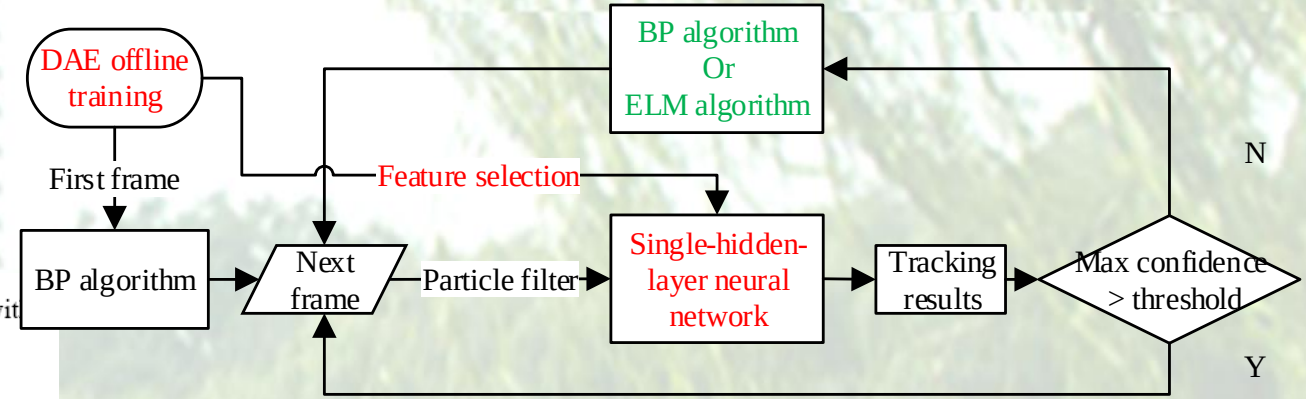
I. INTRODUCTION

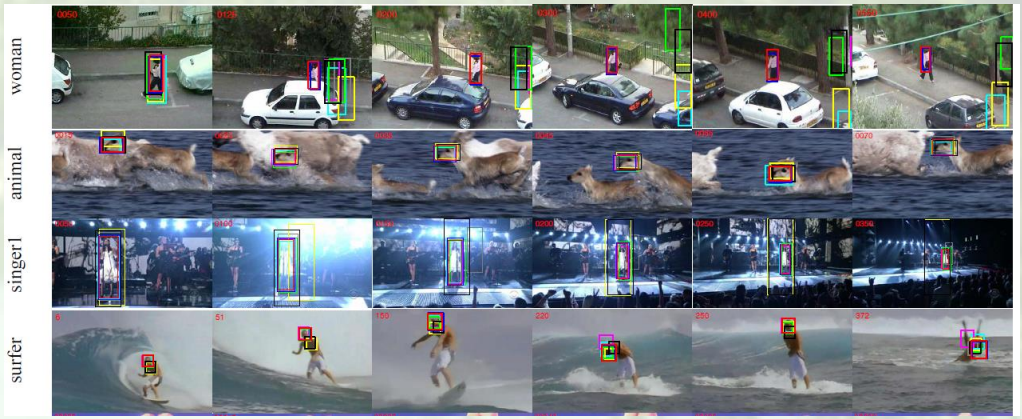
Visual tracking is one of the most important research and development fields in computer vision. However, accurate tracking of general objects in a dynamic environment is still a challenge due to the influence of many factors such as dynamic appearance changes, partial occlusion, cluttered background and distraction from similar objects. In general, the visual tracking techniques can be categorized as discriminative or generative methods. The former formulates the tracking as a classification problem in which the target is discriminated from background and the classifier is updated online during the occlusion and variations since they explicitly take the background into consideration.

Recently, a novel tracker, DLT [4], based on deep learning was proposed for robust visual tracking in the complex environment. In the DLT, the principles of generative and discriminative trackers are combined by developing a robust discriminative tracker. Compared with other discriminative trackers, DLT makes full use of stacked denoising autoencoder [5] to learn (unsupervised) generic robust image features from a very large image dataset and then transfers the learnt features to the tracking task. However, it is noted that the deep nets of DLT causes expensive computational costs.

Recent research [7] on the voice recognition task has also shown that shallow networks can learn the complex functions previously learned by deep nets and achieve accuracies previously only achievable with deep models. Thus, we think that single hidden layer neural networks can learn expressive image features for visual tracking. In this paper, one hidden layer neural network is proposed for fast and robust visual tracking based on that the single hidden layer network with a finite number of hidden neurons is a universal approximator [6]. In the proposed approach, a denoising autoencoder (DAE) is used to learn generic image features offline from a large image dataset and then the learned features are transferred to the tracking process. During the online tracking process, the features and classifier layer can be further tuned to adapt to specific object. Because the neural networks contain both linear and nonlinear transformations, the image representations

Deep Learning Tracker Framework





Special Issue Paper

Mathematical Methods in the Applied Sciences

Received 1 April 2011

Published online 15 August 2012 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/mma.2643
MOS subject classification: 97P99; 97R99

A new robust visual tracking algorithm based on Transfer Adaptive Boosting

Songtao Wu, Yuesheng Zhu*† and Qing Zhang

Communicated by T. Qian

A new robust visual tracking algorithm, based on Transfer Adaptive Boosting, is proposed in this paper . The algorithm is able to produce a higher accuracy hypothesis for more effective and adaptive tracking by using certain useful historical instances of the video. The theoretical upper bound for the classification error of the samples and its corresponding theorem are also given. The theoretic analysis and simulation results show that the new method achieves a better performance compared with other methods in terms of avoiding drifting and miss-tracking, even with complex variations and alterations of the object's original appearance, such as occlusions, illuminations, and shape deformations. Copyright © 2012 John Wiley & Sons, Ltd.

Keywords: visual tracking; historical instances; transfer AdaBoost

1. Introduction

Visual tracking is an important component in practical computer vision applications, such as human behavior recognition, human–computer interactions, and security and intelligent surveillance. The challenges of designing a robust tracking algorithm include dealing with the complex backgrounds, occlusions, variations of target appearances, and illuminations.

Recently, there are many research efforts focused on finding effective and efficient tracking algorithms [1–3]. Treating tracking as a binary classification problem is one of most attractive research points [4–6]. The basic idea of this approach is to use some selected positive samples and the negative samples around them to generate a hypothesis that is utilized in the next frame to discriminate the target from its background with an exhaustive search. After the corresponding location of the target is achieved, several image patches are chosen as training samples to generate a new hypothesis for the next frame, and this process is iterated through all of the following frames. The key is how to effectively choose training samples. A typical method is based on the online AdaBoost (OAB) framework [4] in which some of the simple features, such as local binary pattern, histogram of gradient, or haar-like feature, are selected to form a strong classifier to represent the target. Only the detected image patch is chosen as positive sample and other four patches around it as negative samples in the OAB method. However, the error in each frame introduced in online adaptation process would be accumulated and finally result in drifting. Instead of updating weak classifiers from the found samples, online semi-boost [5] uses a prior model to determine the labels of chosen samples that are utilized to update all the weak classifiers. Once the target suffers occlusions, or any other changes that result in the loss of tracking, the prior model would pull back the tracker to track the target again. This approach may alleviate the drift, but it is nonadaptive and extremely unstable. In literature [6], a tracker based on multiple instance learning (MIL) is proposed to overcome drift. It is different from previous methods in that it directly deals with rural instances; a bag that contains several instances is used to generate a robust classifier. Even if the location found is not precise enough, the tracker can still keep tracking the target on the basis of a MIL framework. Nevertheless, the high complexity of MIL method limits its applications. On the basis of the assumption that the distribution of sample labels obeys some structural constraints, the samples of image patches close to the target trajectory should be positive and those surrounding to the trajectory are negative, an efficient tracker with P-N learning [7] is developed. However, the positive and negative (P-N) constraints cannot pledge the correctness of bootstrapped samples and may lead to miss-tracking. A boosting-based tracker with cotraining [8] uses a boosting error bound in the cotraining framework to guide the tracker. This method assumes that two views are independent but it may not be suitable in most cases. In the method of literature [9] the target is segmented into several non-overlapping parts and each part is tracked simultaneously, but error accumulation is still an unsolved problem.

In this paper, a new robust visual tracking algorithm based on the Transfer Adaptive Learning [10] algorithm is proposed to improve traditional boosting-based tracking algorithms. Unlike the methods mentioned earlier, in which only the samples in the current frame

RAPID VISUAL TRACKING WITH MODIFIED ON-LINE BOOSTING AND TEMPLATE MATCHING

Guibo Luo, Yuesheng Zhu, Qing Zhang
Communication & Information Security Lab
Shenzhen Graduate School
Peking University, Shenzhen, China
luoguibo@sz.pku.edu.cn, zhuys@pkusz.edu.cn, fjzhangqing@gmail.com

ABSTRACT

On-line learning is increasingly popular in visual tracking, but the challenge that it faced is how to adapt the appearance changes and avoid the drifting or missing track. In this paper, a fast visual tracking algorithm is proposed to make the tracker more accurate and stable in the complex variations situations like occlusions, illuminations and shape deformations. In the proposed algorithm, a modified on-line boosting method is developed to make the tracker more adaptive to variable scene and a template matching model is used to constrain the training samples, so that the accumulating errors in self-update learning can be alleviated effectively. In addition, an optimization process is used to reduce the computational burden. Our experimental results have demonstrated that compared with other on-line tracking (MIL) algorithm to reduce the label jitter phenomenon during updates in the tracking. The basic idea of MIL paradigm is that during the training process the samples are packaged in bags rather than individual samples, that is, instances close to the current object position are considered as a positive bag and other instances as negative bags, and the classifiers are trained to discriminate these bags instead of samples. Kalal et al. [6][7][8] exploited a parallel framework including tracking, learning and detection (TLD) for long-term tracking. TLD uses a Lukas-Kanads-tracker(LKT) to guide the learning process of the object detector. This selective-update strategy can achieve promising results in long-term tracking. However, the LKT is not competent to track the objects with non-rigid transformations and partial occlusions

Research Article

Mathematical Methods in the Applied Sciences

Received XXXX

(www.interscience.wiley.com) DOI: 10.1002/sim.0000
MOS subject classification: XXX; XXX

A Robust Tracking Method with Adaptive Local Spatial Sparse Representation

Qing Zhang, Yuesheng Zhu*, Songtao Wu, Guibo Luo, and Liming Zhang

In this paper, based on local spatial sparse representation (LSSR) a robust visual tracking method is proposed. In the proposed approach, the learned target template is expressed sparsely and compactly by forming local spatial and trivial samples dynamically, an adaptive multiple subspaces appearance model is developed to describe the target appearance and construct the candidate target templates during tracking, and then an effective selection strategy is employed to pick out an optimal sparse solution and locate the target accurately in the next frame. The experimental results have demonstrated that our method can performs well in the complex and noisy visual environment, such as heavy occlusions, dramatic illumination changes and large pose variations in video Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: Visual tracking, sparse representation, trivial samples, multiple subspaces

1. Introduction

視頻智能分析實例

- 應用1：視頻搜索系統
- 應用2：視頻跟蹤系統

北京大學

圖像智能分析實例（Demo）

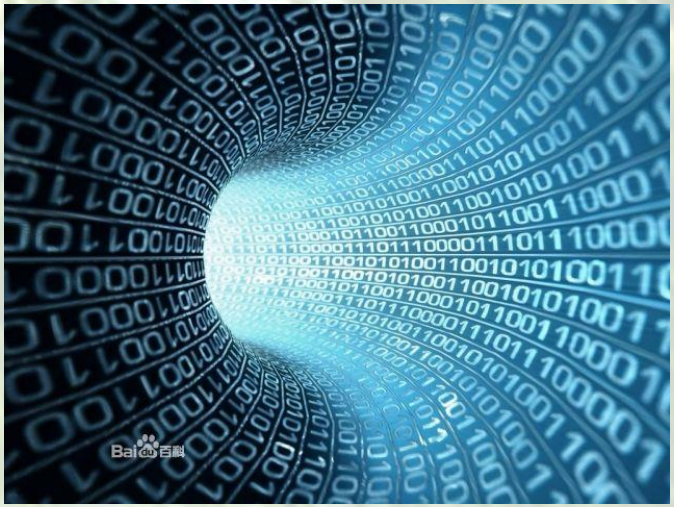
基于数字指纹的图像管理系统

基于数字指纹的图像管理系统：图像列表

北京大学

内容提要

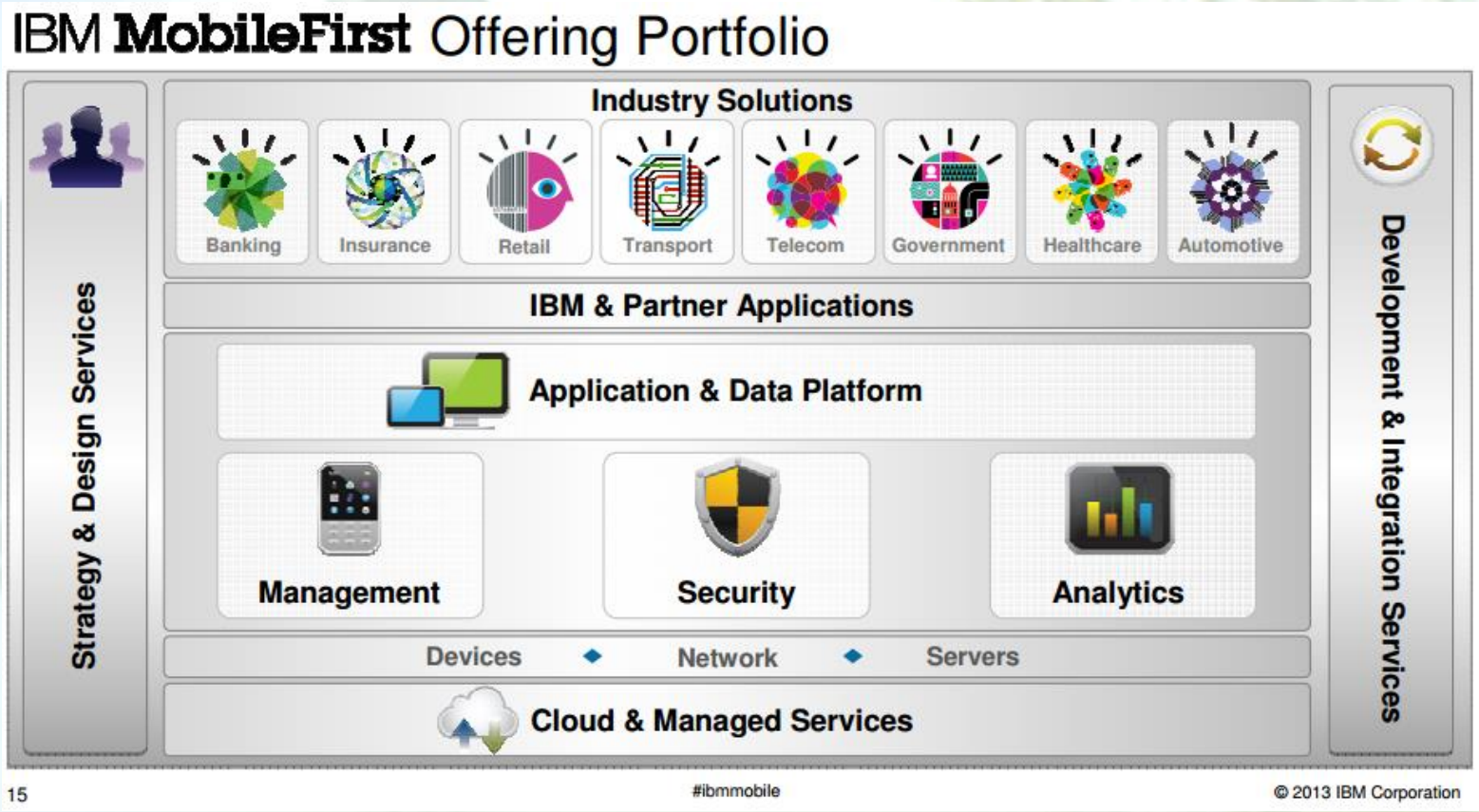
- 大數據基本概念
- 大數據理論基礎
- 非結構化大數據智能分析技術
- **大數據的安全問題**



北京大学

Apple硬體+IBM軟體

IBM為Apple提供應用解決方案：
將大數據分析和安全技術（**MobileFirst**）引入**Apple**移動設備



“大數據”的安全問題

所有數據都被記錄和保留下來，被分析加工和利用

- **政府的敏感數據**
  被濫用將成為影響社會穩定的安全隱患
- **企業敏感數據**
  被濫用將成為利益衝突及商業糾紛的安全隱患
- **個人隱私/敏感數據**
  大量的數據匯集囊括了個人隱私，以及各種行為的細節記錄
  行為分析/被濫用將成為危及人身安全的隱患。
- **數據安全管理**
  破壞、丟失、盜取

- 索尼的系統漏洞導致7700萬使用者資料失竊
- iOS被發現按照時間順序記錄使用者的位置座標信息

北京大學

挑戰：數據採集 VS 信息安全

挖掘利用將可侵犯個人隱私

矛盾？

關鍵點：

- **誰在取？**
- **從哪取？**
- **放在哪？**
- **誰在管？**
- **誰在挖？**
- **挖給誰？**
- **誰在用？**

網路與信息安全技術

大數據安全與隱私保護關鍵技術

數據發佈匿名保護技術
社交網路匿名保護技術
數據指紋/水印技術
數據溯源技術
角色挖掘
風險自我調整的存取控制

進不來
拿不走
看不懂
改不了
賴不掉

道高一尺，魔高一丈的博弈

北京大學

抗惡意攻擊的秘密共用方案及應用

- 將秘密以適當的方式拆分，分割後的每一個份額由不同的參與者管理

  - **秘密K 被拆分為n 個份額**
  - **任意 t（t≤n）個或更多個份額可以恢復秘密 K**
  - **任何t – 1或更少的份額恢復K**

單個參與者無法恢復K，只有若干個參與者一同協作才能恢復K

- 目的
  避免秘密過於集中，達到分散風險和容忍入侵的目的

- 应用
  分布式網路環境中保護數據安全。在信息化和網路管理及控制有廣泛應用

北京大學

Q&A

北京大學信息工程學院教授/博導
北京大學大數據技術研究院
通信與信息安全實驗室主任
深圳寬頻無線網絡安全技術工程實驗室主任

中國大數據專家委員會委員
深圳雲計算產業協會副會長

聯繫方式

Email: zhuys@pku.edu.cn

北京大學