Predictive to Prescriptive Analytics
Challenges & Opportunities

By:

Prof. Dr. O.P. Vyas
Professor (Information Technology)
Program Mentor (Business Analytics)

Indian Institute of Information Technology Allahabad
Outline of Talk

• Brief about Analytics Approaches
  Research Interest & Projects
  • Research group (Data Science)
    • Research Work & Collaborations
    • Indo-German Collaboration

• Business Data Analytics Approach
  • Data Analytics: Descriptive, Predictive & Prescriptive- Overview
  • Business Intelligence & Prescriptive Analytic approach: Recent trends
  • Prescriptive Analytics: Overview & Applications

• Prescriptive Analytic approach : Critical Infrastructure in Smart Cities
  • Smart City mission of Govt. of India: What makes a city smart...?
  • CIIs & Vulnerabilities in the context of Smart Cities
  • CPSEC: Cyber Physical Security of Energy Infrastructure in Smart Cities
  • Cyber Security: Attack Vectors & Trend Analysis

• Cyber Security Framework
  • Existing cooperation & Proposed Approach
International Academic Cooperation & Research Projects

- Broadly working in the fields of **Data Science**, the Research Group had following collaborations;
  1. **Indo-German** Cooperation during 2007-19.
     
     Having completed *DAAD funded* academic stay from July 1995 to Sept. 1997 resulting into dual-degree of *M.Sc. nat. from TU Kaiserslautern* (Germany) and consequently *Ph.D. from IIT Kharagpur in 2000*, the Research Group headed by Prof. O.P.Vyas had extensive academic cooperation with German Univ./Institutes.
     
     a) Technical University of Kaiserslautern.
     b) Fraunhofer Institute of Software Engineering, Kaiserslautern.
     c) Fraunhofer Institute for Autonomous Intelligent Systems, St. Augustin-Bonn.
     d) Rostock University.
     e) University of Paderborn.
     f) Bielefeld University.


  3. **Indo-French** cooperation project ‘*IoT deployment and service provisioning for Smarter cities*’ with Dr. Nathalie Mitton of Inria-Lille, France during 2014-15.

     ✓ Each side visiting for collaborative work with students internships

  4. **Indo-Italian** collaboration- “*Data Analytics in Smart city environment*” with Univ. of Messina, Italy. 2014-19 SmartME Project.

     ✓ Academic exchange involving Faculty members, Research Scholars and Students with Master / Bachelor thesis in collaborative work.

Indo-Germany: Academic Cooperation & Projects

- The Research Group headed by Prof. O.P. Vyas had academic cooperation for over a decade with German Univ.
  - Successful completion of joint Ph.D. supervision with Prof. Stefan Wrobel of Bonn Univ. (Fraunhofer IAIS - St. Augustin) during 2008-09.
    - Joint research supervision-Indian PhD. Scholar Mr. Lokesh Sharma spent 02 years at Bonn under DAAD Sandwich model Ph.D. scheme. ‘Spatio-Temporal Data Mining: Homogenization of Clusters’ was central theme of research work.
  - BMBF (Germany) & DST (India) funded Research project ‘Click Stream mining for Web content syndication’ with Prof. Clemens Cap-Rostock University Germany during 2007-09.
    - Effected Academic exchange of Academic personnel for collaborative project work.
  - Max Planck Society funded Project “Service Oriented Architectures” with TU Kaiserslautern (Germany) 2009-2010.
    - Academic exchange as Dr. Bernd Reuther visited India and ~04 researchers from India visited Germany for collaborative work.
    - Extensive cooperation with joint PhD supervision, Teaching of short courses with 04 visits each of Faculty members from both sides and ~10 students internships at Uni Paderborn.
    - Prof. Artus Krohn-Grimberghe is appointed as Adjunct Professor at IIIT Allahabad in 2018.
IIIT-A & German Univ.: Collaboration….

- Being Alumnus of a German University TU Kaiserslautern and having working experience as Guest Professor at Uni Paderborn for past 5 years, I can see many mutually beneficial possibilities with Germany & India, specially when IIITA is now in readiness for Internationalization.

- India and Germany has history of long successful associations:
  - Both the Higher Education system has complementary and mutually beneficial skills and strategic advantages in various disciplines of Higher education.
  - Various funding agencies like DAAD, Humboldt foundation, BMBF, DFG etc from Germany and UGC, MHRD, DST from India providing structured framework under Indo German cooperation since past many decades.
    - From Bachelor level, Master level and Doctoral students to Faculty exchange with Project funding.
  - Readiness of highly motivated and talented students to undetake assignments with German Academics.
Data driven Intelligence…

- **Curate**: Collect, store, manage, clean, process data
- **Summarize**: Provide tables, charts and other summaries
- **Describe**: Identify meaningful patterns in data
- **Predict**: Understand causal relationships to predict future scenarios
- **Prescribe**: Create the future through informed, value-driven choices

**Analytics**: Descriptive / Predictive / Prescriptive
Data Analytic Approaches as Descriptive & Predictive

Data Analytics

- Descriptive
  - Association Rule Mining
  - Clustering Analysis

- Predictive
  - Classification Techniques
  - Associative Classifiers
Data Analytic Approaches as Descriptive & Predictive

- **Data Analytics**
  - **Descriptive**
    - Association Rule Mining
    - Clustering Analysis
  - **Predictive**
    - Classification Techniques
    - Associative Classifiers
Association Rule Mining is one of the most common and useful types of data mining.

- The purpose is to determine the “interdependence” among various items.

Support-Confidence Framework

- **Apriori** (Agarwal et.al. 1993)
- **Sequential pattern mining** (Agarwal et.al.’95)
- **Hashing in Apriori** (Park et.al. 1995a)
- **FP Growth** (Han et.al. 2000): ARM without Candidate generation.
Clustering is the process of grouping a set of (physical or abstract) objects into classes of similar objects.

**Spatial Data Mining & Web Mining**
- K-Means Algo (Mac-Queen et al., 1967)
- CLARANS (Han et al., 1994)
- DBSCAN (Ester et al, 1996)

**Search Engine Result Mining**
- Clustering Search Result (Leouski and Croft, 1996, Zamir and Etzioni, 1997).
- PageRank Algorithm – Google 1996
Data Analytics

Descriptive
- Association Rule Mining
- Clustering

Predictive
- Classification mining analyzes a set of training data (i.e. a set of objects whose class labels are known) and constructs a model for each class based on the features in the data. A set of classification rules are generated by the classification process, and these can be used to classify future data, as well as develop a better understanding of each class in the database.
- Associative Classification
Prescriptive Analytics: Paradigm shift or passing fad...

The paradigm shift in Information delivery was effected with Prescriptive Analytics, which is also being referred as third and final frontier in Analytics....
Significant Issues in Prescriptive Analytics approach

➢ How does the market define prescriptive analytics?
➢ What is the difference between prescriptive analytics and predictive analytics?
➢ Is prescriptive analytics right for your business, and are you ready?
➢ Different approaches to prescriptive analytics, its transformational value.
➢ How prescriptive fits in the emerging trends like AI and ML.
➢ Opportunities.
➢ Challenges.
What is Prescriptive Analytics?

“Prescriptive analytics is the application of logic and mathematics to data to specify a preferred course of action. While all types of analytics ultimately support better decision making, prescriptive analytics outputs a decision rather than a report, statistic, probability or estimate of future outcomes.”

Prescriptive Analytics

- **Prescriptive analytics** being referred as the *third and final* phase of Business analytics also includes Descriptive and Predictive analytics.
- Prescriptive Analytics is considered as the next step towards increasing data analytics maturity and leading to *optimized decision making*, ahead of time.
  - Compared to **Descriptive** and **Predictive** Analytics, the Prescriptive approaches are rather less mature but attracting increasing research interests.
  - It aims at suggesting (**prescribing**) the best decision options in order to take advantage of the predicted future.
    - To do this, it incorporates the **predictive analytics output** and utilizes **optimization algorithms in probabilistic context** to provide **adaptive, automated, constrained, time-dependent** and **optimal decisions**.
From **Descriptive to Prescriptive Analytics**

- The Analytic approaches though answer different Questions but more often;
  - Descriptive Analytics naturally leads to Diagnostic Analytics.
    - What happened can naturally take you to Why it happened.
  - Predictive Analytics naturally leads to Prescriptive Analytics.
    - What will happen can naturally take you to How can we make it happen.
- **Prescriptive Analytics** though has some bonafide elements to it.
The Objective is to explore as to how to apply prescriptive analytics as a form of decision support for enabling business houses to answer their most pressing problems.
If you search Google trends today, you’ll see that interest in the topic “prescriptive analytics” has grown significantly since Google began collecting this data in 2004. The real boom began in 2013, and a rapid growth is observed in interest since then. It’s worth noting that there is no sign of the trend curve flattening out, as it continues to grow each year.

Interest in prescriptive analytics from 2004 to today
Operations Research in Prescriptive Analytics
A BRIEF INTRODUCTION TO PRESCRIPTIVE ANALYTICS

The major mathematical-based disciplines of prescriptive analytics include:

- Operations Research
- Machine Learning
- Natural Language Processing
- Applied Statistics

Research firms, vendors, consultants, and market leaders have trended toward dividing prescriptive analytics into two different approaches:

- Heuristics-based automated decision making
- Optimization-based decision support
HEURISTICS-BASED DECISION AUTOMATION

- Rules-based decision automation is different. It means that when something happens, the system will decide what to do on the fly, given a set of predefined rules that have been plugged in.

- Unlike optimization, this approach cannot provide an answer outside of what has been predetermined. Further, the “math” is very different from the math used in optimization. It typically uses a form of statistics and applies algorithms to find the answer.

- Heuristic algorithms do not guarantee the best answer. If designed well, they can offer a short-cut approach to finding good answers in a reasonable amount of time.

So it might be possible that Exact algorithms guarantee the best answer. However, for difficult problems, the time to solve for the best answer can increase exponentially compared to the size of the problem.
➢ To solve operational problems, such as route optimization and logistics planning, Operations Research professionals traditionally applied optimization.

➢ Advanced optimization models combine the value chain (including key constraints) with financials, providing higher quality information than what’s possible with single predictive or BI models.

➢ Applying prescriptive analytics through optimization enables users to wade through all these factors and find the path that meets the most objectives given the defined business.

➢ So it uses the mathematical algorithms that maximize or minimize one or more objective functions while still respecting business realities, thus always producing feasible plans.
VALUE ADDITION OF PRESCRIPTIVE ANALYTICS TO BUSINESSES

➢ ACHIEVE HIGHER CONFIDENCE IN PLANS PLUS LOWER RISK

➢ ESTABLISH HIGHER AGILITY IN THE ORGANIZATION

➢ MITIGATE RISK

➢ EARN A HIGHER RETURN ON EXISTING ASSETS

➢ ADDRESS NEW PLANNING CHALLENGES USING THE BEST METHOD POSSIBLE

➢ IMPROVE PERFORMANCE
  ➢ Improving the effectiveness of the business against one or more objectives (i.e., operating income, net income)
  ➢ Increasing the efficiency of an operation (i.e., do more with same resources, achieve the current outcome with fewer resources)
  ➢ Maximizing the return from altering the design of a system, subject to a defined maximum risk.
"Companies like PepsiCo are all working on moving from the traditional descriptive and diagnostic analytic capabilities to prescriptive analytics. Adopting prescriptive is critical for supply chains to gain a competitive advantage now and in the future."
– Leslie Keating, Former SVP Supply Chain, Frito-Lay

Next Slide adduces the opportunity present in different industry to use Prescriptive Analytics to enhance their business outcomes.
<table>
<thead>
<tr>
<th>INDUSTRY</th>
<th>SOLUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial Services</td>
<td>Cash Management, Mortgage Services, Strategy, Portfolio Optimization</td>
</tr>
<tr>
<td>High Tech</td>
<td>Integrated Business Planning</td>
</tr>
<tr>
<td>Aerospace &amp; Defense</td>
<td>Service Contract Profitability Modeling</td>
</tr>
<tr>
<td>Healthcare (Providers)</td>
<td>Health Plan Benefit Design Optimization Staff, Service and Resource Optimization</td>
</tr>
<tr>
<td>Utilities</td>
<td>Operational Planning (weekly planning to 25+ year long-range planning)</td>
</tr>
<tr>
<td>Consumer Packaged Goods</td>
<td>Trade Promotion Optimization, IBP/Sales and Operations Planning (S&amp;OP)</td>
</tr>
<tr>
<td>Oil &amp; Gas</td>
<td>Logistics Optimization, Commodity Trading Optimization</td>
</tr>
<tr>
<td>Retail</td>
<td>Price and Promotion Optimization</td>
</tr>
<tr>
<td>Chemicals</td>
<td>IBP/S&amp;OP CAPEX</td>
</tr>
<tr>
<td>Government</td>
<td>Army Recruiting</td>
</tr>
<tr>
<td>Natural Resources</td>
<td>Network Optimization, IBP/S&amp;OP CAPEX</td>
</tr>
<tr>
<td>Metals</td>
<td>Product Mix and Supply Planning</td>
</tr>
<tr>
<td>Mining</td>
<td>Supply Chain Planning, Blend Optimization</td>
</tr>
<tr>
<td>Public Sector</td>
<td>Personnel Training Optimization</td>
</tr>
<tr>
<td>Telecomm-unications</td>
<td>Optimization of Channel Allocation to Spectrum</td>
</tr>
</tbody>
</table>

Real-Life, Cross-Industry Applications of Prescriptive Analytics
Three Phases in Prescriptive Analytics
Predictive & Prescriptive Analytics

- Although a Predictive analysis itself can also lead to a prescription,
  - “If the stock prices are predicted to go up in next days, it is recommended to go for Buy”

- Though the Question behind this Prediction is an optimization problem (optimizing the profit) but the optimization itself is not the part of Predictive Analytics.

- Prescriptive Analytics on the other hand consists of answering the Questions related to “What if” situations for single or multiple Predictions with optimality condition as part of it. Many researchers consider following two parts as essential for Prescriptive Analytics;
  - Exploration of possible events
  - Generation of the Prescription

- Decision space for Prescriptive Analytics tends to be larger; multiple situations with many variables, options and constraints taken into account
Categories of Predictive Analytics Methods

- Probabilistic Models
  - Bayesian Network
  - Markov Chain Monte Carlo
  - Hidden Markov Model
  - Pattern recognition
  - Random Forest
  - Gaussian process
  - Conditional inference tree
  - Support Vector Machine
  - Ensemble learning
  - Artificial Neural Network
  - Random search
  - Decision tree
  - Clustering-based heuristics
  - k-nearest neighbors algorithm
  - Kernel methods
  - Multilayer perceptron
  - Gradient Boosted Tree

- Machine Learning / Data Mining
  - Linear regression
  - Multiple linear regression
  - Rank regression
  - ARIMA
  - Logistic regression
  - Multinomial logistic regression
  - Density estimation
  - Support vector regression

- Statistical Analysis
Categories of Prescriptive Analytics Methods

- Probabilistic Models
  - Markov Decision Process
  - Hidden Markov Model
  - Markov Chain
  - K-means clustering
  - Reinforcement Learning
  - Privacy preservation
  - Boltzmann Machine
  - Nadaraya-Watson estimator
  - Artificial Neural Networks
  - Mixed Integer Program
  - Linear Program
  - Binary Quadratic Program
  - Non-Linear Program
  - Binary Linear Integer Program
  - Stochastic Optim
  - Conditional Stochastic Optim
  - Constrained Bayesian Optim
  - Fuzzy Linear Optim
  - Robust and Adaptive Optim
  - Dynamic Program
  - Optimal searcher path
  - Genetic algorithm
  - Evolutionary Optim
  - Greedy algorithm
  - Particle Swarm Optim
- Machine Learning / Data Mining
  - Simulation over Random Forest
  - Risk Assessment
  - Stochastic simulation
  - What-if scenarios
- Mathematical Programming
  - Association rules
  - Decision rules
  - Criteria-based rules
  - Fuzzy rules
  - Distributed rules
  - Benchmark rules
  - Desirability function
  - Graph-based recommendation
- Evolutionary Computation
- Simulation
- Logic-based Models
From Predictive to Prescriptive Analytics

As Prescriptive Analytics takes as input the outcome of Predictive Analytics algorithms, more closer look at various Predictive approaches can see following three major categories of Predictive Analytics approaches;

1. Probabilistic Models
2. Machine Learning / Data Mining
3. Statistical Analysis
Prescriptive Analytics

Prescriptive Analytics is to suggest actions to decision maker based on descriptive and predictive analytics.

Steps:
❖ Frame the problem
❖ Identify candidate actions
❖ Predict consequences of actions
❖ Assess value of consequences
❖ Suggest highest value actions

Common categories of analytical methods
❖ Optimization
❖ Simulation
Prescriptive Analytics: Optimization, Simulation

● **Key task**: Create a model – A person (you) must do this
  ○ Model must capture **essential features** of the business situation
  ○ Larger models often get their data from BI / Descriptive Analytics

● **Given an appropriate model, we can**:  
  ○ Ask “*What are all the possible outcomes?*” – **simulation/risk analysis**  
  ○ Ask “*What’s the best outcome we can achieve?*” – **optimization**
Prescriptive Analytics Framework

In traditional **Predictive Analytics** we seek to model past data $Y$ using a **model M**. The distribution $P(y|M)$ produced by such a model can be used as a predictor for future values of $y$.

For example, suppose $Y$ captures part failures over time in a population of vehicles in the field. Part failures are usually modeled using a Weibull distribution, using which the expected number of failures in the future, say $y_e$, can be computed using $P(y|M)$. (*The prediction procedure in Figure is formulated as full Bayesian model selection, i.e., the ‘most probable’ model $M$ is chosen, given the available data $Y_1$.*)

The PAF is also called as Open-loop **Prescriptive Analytics Framework**
There are cases when feedback to the system improves the results, as when the optimal strategy $x$ is put into practice, it becomes part of the data.

For example, by modeling monthly sales data $Y$, the expected future sales can be forecasted using a statistical model. More interestingly however, a predictive model resulting a distribution $P(y|x, M)$ can be used to estimate one component of $y$, say sales quantity, from other elements such as demographics, etc. as well as the quoted price $x$. 

Fig. : Adaptive Prescriptive Analytics
Prescriptive Analytics

Theoretical Foundations: Optimization based
From its foundation Machine Learning has largely focused on Supervised Learning or the prediction of a quantity $Y$ (usually univariate) as a function of $X$ based on Data $\{(x_1, y_1), \ldots, (x_N, y_N)\}$. (Training data as well as testing)

Though there are large number of predictive techniques with varying performance but ML however does not address optimal decision making.

Some Researchers have proposed to combine ideas from Machine Learning (ML) and Operations Research and Management Science (OR/MS) in developing a framework, along with specific methods, for using data to prescribe optimal decisions in OR/MS problems.
Optimization in Prescriptive Analytics

Some seminal research contribution highlights the issues with the help of stochastic optimization problem while formulating a predictive prescription that prescribes a decision in anticipation of the future.

In the proposed work the problem of prescribing an optimal decision is addressed in a framework where its cost depends on uncertain problem parameters $Y$ that need to be learned from data.
Budget optimal ≠ cost optimal
Let's look at the optimization problem as below, which have been proposed by researchers to formulate Prescriptive Analytics.

Given a variable $Y$ (Traffic), we have to optimize the loss function $L(z,Y)$, which is taking arguments as set of all possible decision $z$ with $Y$ as Training data.

$L$ is typically a Loss function that has be minimum, $Z$ is the decision constrained in $E^{z*} = \arg\min_{z \in Z} \mathbb{E}[L(z,Y)]$ made after some observation $X = x$ with the objective of minimizing the uncertain costs $L(z,Y)$ and $\mathbb{E}$ represents the expected value from some random variable.

Given this context our objective is to find the $z^*$ that will be the best optimal action among all feasible candidate.
Observations

➢ Prescriptive analytics is still a work in progress with many different formulations and in general, Prescriptive analytics seeks to determine the best solution or outcome among various choices.
   ➢ But the three branches of analytics can’t be separated.
   ➢ There is always need to do descriptive and predictive before jumping into prescriptive.

➢ The Bootstrap Prescriptive Analytics proposed by Dimitirs et al is a proactive analytics technique where a continuous improvement in the model (Predictive Analytics) is done in order to achieve the best optimal prescription.
Prescriptive Approach in Smart City context
Protecting Critical Infrastructure: Context Smart Cities

Prescriptive Analytic Approach
Government of India has announced an ambitious 100 Smart Cities mission.

State capitals, and many tourist, heritage cities are expected to witness a rapid upgrade of urban infrastructure and online services to citizens, enabled by Information Technology.
A smart city is an urban development vision to integrate multiple ICT solutions in an efficient fashion to manage a city’s assets for its citizens.

The goal of building a smart city is to improve quality of life by using technology to improve the efficiency of services and meet residents’ needs.
What makes a Smart City (Building, Transportation, Energy System)?

- **Monitor**: Sense/Observe, Communicate/Connect
- **Model**: Analyze/Optimize
- **Manage**: Present/Visualize, Act

Smart Cities is termed *Smart* because of its *capacity to learn from past* and take better *decision in autonomous* manner.…

Data *collection and Analysis* to affect the system
Act in *Real-time @ City Scale*
Drives *Learning algorithms* and processing architecture

Ref: Randy H. Katz, The Software-Defined Building a better machine for living, 10th Intl Conference on Network and Service Management, Rio de Janeiro, Brazil 19 November 2014
Every Smart City Model is characterized by a Central Command and Control Centre, which has all the Critical Resources and Infrastructure monitored, managed and controlled at one Place!!

Should we not be sufficiently alerted with this as a potential Threat!!
City Command Centre of Ahmedabad-India
City Command & Control Centre: Smart City

- Smart Transport
  - Traffic Mgmt
  - Parking
  - Area Traffic Control
  - Road Condition System
  - Real Time Travel Response

- Smart Streets
  - Digital Signage
  - Streetscape

- Smart Landscape
  - Irrigation System

- Smart Desks
  - City Information
  - Facilitation Centers

- Smart Buildings
  - Offices
  - Home
  - Hotel
  - School
  - Hospital

- Smart Urban Infrastructure
  - Water Systems
  - Power
  - ICT System
  - District Cooling
  - Domestic Gas
  - Waste Management
  - Fire Fighting
  - Utility Tunnel

City Command and Control Centre (C4)

Smart City Security and Surveillance
Smart Cities Technology Centre

Intelligent Urban and Environmental Analytics and Systems
- Optimization
- Predictive Modelling
- Forecasting
- Simulation

Driving New Economic Models
Significant Collaborative R&D
Skills Development & Growth
Competitive Advantage

Collaboration and Access to Local, Regional & Worldwide Network
SME’s | MNC’s | Universities | Public Sector | VC Community

Instrumented
City Fabric
Energy
Movement
Water

Seed Projects
Real World Insight | Data Sets | Devices
As advancements in ICT infrastructure grow our vulnerability to damage through attacks by insurgents/terrorists with objective to immobilize and paralyze day-to-day activities of the nation is becoming real.

Such damage would cause short and long term setback to economy. We have many lessons from Global initiative to secure our cyber system, while planning and implementing India’s ICT infrastructure.

Natural or insurgency/terrorist induced disaster increases pressure on available ICT systems exponentially to facilitate coordination between various agencies to protect Critical Infrastructure ?.
Critical Infrastructure

Systems, Networks and Assets that are so essential that their continued operation is required to ensure the security of a given nation, its economy, and the public's health and/or safety

- Banking & Finance
- Transport (Rail, Road, Air)
- Industrial Production
- Water Resources / Dams
- Govt. Services (Health, Emergencies)
- ICT Infra

Any Country depends completely on various systems such as Power Grid, ICT Infrastructure, Transport system as any disturbance creates cascading effect in the Cyber Attack..
Why **Cyber Security in CI** more pertinent now..

Computerised equipments are used in the control of equipment and industrialised processes and these are deployed in every aspect of **Critical National Infrastructure**, such as Nuclear Power Plants; **Electric Grid**, Chemical Plants; Oil Refineries; Gas processing Facilities; **Railway and Transport Communication systems**; Pharmaceutical productions; LPG tankers; **Distribution Centres and Ports**.

They are **Cyber-Physical Systems** ..
Critical Infrastructure

Most of the Infra Power Grid, Nuclear Plant, Industrial Automation are Cyber-Physical systems

- High Interdependence in them...
  - If Power goes down, Internet does not work so the related effect...
  - If Nuclear Plant System can be manipulated, then catastrophe can not be avoided..

- High Increase in Automation and Centralized Control....so the centralized destruction is also becoming easy, possible..?

- Known Risks and Issues
  - What are the Security - Prevention mechanism for these issues?

- What are the immediate next steps?
  - Provincial, National and Global Level.
Cyberspace & physical space are increasingly intertwined and software controlled/enabled

Need for secure software applications

Agriculture and Food  Energy  Transportation  Chemical Industry  Postal and Shipping
Water  Public Health  Telecommunications  Banking and Finance  Key Assets

Critical Infrastructure / Key Resources
- Farms
- Food Processing Plants
- Power Plants
- Production Sites
- Railroad Tracks
- Highways
- Bridges
- Pipelines
- Ports
- Chemical Plants
- Delivery Sites
- Nuclear Power Plants
- Government facilities
- dams

Physical Infrastructure
- Reservoirs
- Treatment Plants
- Hospitals
- Cable Fiber
- FDIC Institutions

Cyber Infrastructure
- Internet
  - Domain Name System
  - Web Hosting
- Control Systems
  - SCADA
  - PCS
  - DCS
- Services
  - Managed Security
  - Information Services
- Hardware
  - Database Servers
  - Networking Equipment
- Software
  - Financial System
  - Human Resources

“In an era riddled with asymmetric cyber attacks, claims about system reliability, integrity and safety must also include provisions for built-in security of the enabling software.”
Energy Infrastructure : Cyber Physical threats

The **Energy Production and Distribution systems**, which comprise a vital economic and social infrastructure, are exposed to:

(a) Security threats inherited from the ICT sector, **Cyber threat**  
(b) Physical attacks like bombing, fires and floods, and  
(c) Combinations of cyber and physical threats.

Considering that a potential attack to an energy installation may lead to **cascading failures**, these threats can lead to disruptions with severe consequences like destruction of other interconnected CIs (**e.g., water, communications and transportation**), loss of human lives and environmental impact.
Cyber-Physical Security: Protecting the CI

Proposed Approach
Why Cyber Attackers are successful in reaching Critical Infrastructures

Cyber Attack Points - ICS & SCADA

Supervisory Control And Data Acquisition are the basis of all successful Automation Projects worldwide.

SCADA is a computer-based system for gathering and analyzing real-time data to monitor and control equipment that deals with critical and time-sensitive materials or events.

SCADA systems were first used in the 1960s and are now an integral component in virtually all industrial plant and production facilities

IN EARLY PHASE SCADA WAS ISOLATED BUT NOW CONNECTED TO INTERNET ..!

Vulnerabilities in SCADA have been identified by the Cyber Army...
Power Subsystem Vulnerabilities

• The advantages of digital microprocessors, electric power utilities have exploited computer technology for improved communication and automation of control centers, substations and remote protection equipment. Significant is: Supervisory Control And Data Acquisition (SCADA) systems.

• The use of these computer-based systems for electric power control and protection has created vulnerabilities within the power grid similar to those seen in more traditional computer networks.

Physical intruders have been known to randomly or maliciously push buttons and operate circuit breakers, Reclosers and switches; We must assume that electronic intruders would, too.
# Cyber-Physical Security in Energy Infrastructure of Smart Cities (CPSEC)

## Indo Norwegian Project Proposal

(Under DST India-RCN Norway Joint Call on *Smart Environments 2017*)

<table>
<thead>
<tr>
<th>Investigator (Indian Side)</th>
<th>Investigator (Norwegian Side)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prof. Dr. O. P. Vyas</strong></td>
<td><strong>Prof. Sokratis Kastikas</strong></td>
</tr>
<tr>
<td>Professor (IT)</td>
<td>Professor (IT)</td>
</tr>
<tr>
<td>Department of Information Technology, Indian Institute of Information Technology Allahabad (IIIT-A)*</td>
<td>Center for Cyber and Information Security and Communication Technology (IITK) Norwegian University of Science and Technology (NTNU), Teknologiveien 22, 2815 Gjovik, Norway</td>
</tr>
<tr>
<td>Deoghat, Jhalwa, Allahabad 211012, (UP), India</td>
<td>E-Mail: <a href="mailto:sokratis.kastikas@ntnu.no">sokratis.kastikas@ntnu.no</a></td>
</tr>
<tr>
<td>Email Id: <a href="mailto:opvvas@iitd.ac.in">opvvas@iitd.ac.in</a></td>
<td></td>
</tr>
<tr>
<td><strong>Prof. Sandeep K Shukla</strong></td>
<td><strong>Dr. Pankaj Pandey</strong></td>
</tr>
<tr>
<td>Center for Cyber Security (CSC)</td>
<td>Center for Cyber and Information Security and Communication Technology (IITK) Norwegian University of Science and Technology (NTNU), Teknologiveien 22, 2815 Gjovik, Norway</td>
</tr>
<tr>
<td>Department of Computer Science and Engineering Indian Institute of Technology, Kharagpur (IITK) Kalyanpur, Kharagpur, U.P., India 208016</td>
<td></td>
</tr>
<tr>
<td><strong>Dr. S Venkatesan</strong></td>
<td><strong>Mr. Vikas Katta</strong></td>
</tr>
<tr>
<td>Department of Information Technology IIT Allahabad, Jhalwa, U.P., India-211015</td>
<td>Institute for Energy Technology (IFE) Os Alle 5, NO-1777 Halden, Norway</td>
</tr>
</tbody>
</table>
Four major challenges stemming from the aforementioned issues:

- **C1.** Address cyber, physical and combined cyber-physical threats;
- **C2.** Consider interconnections among CIs (*Critical Infrastructures*), cascading effects and legacy equipment;
- **C3.** Protect the environment, inform the public, as well as coordinate response teams;
- **C4.** Be cost-effective in order to encourage deployment in a global crisis environment.

The main technical output of the project will be the Integrated Security, Safety and Site Management (IS3M) platform which will cover a wide variety of concepts, including, systemic risk management, prevention by design, monitoring and detection, response and mitigation, and information sharing.
The CPSEC IS3M platform will consist of the following main components:

- **SIMA**: Systemic Risk Management
- **PRIDE**: Prevention by Design
- **MADE**: Monitoring and Detection
- **ROAM**: Response and Mitigation
- **IRIS**: Information Sharing
CPSEC IS3M

- The CPSEC IS3M platform is envisaged as an enhanced integrated platform consisting of **PSIM** (Physical Security Information Management System), **SIEM** (Security Information and Event Management) and **CAD** (Computer Aided Dispatch) modules.

- It is proposed to unify the CPSEC disparate physical security devices (like video cameras, access control, intrusion detection sensors), cyber security systems and legacy monitoring and control systems (like SCADA, IoT, and social media) into one converged platform for security management.

- Moreover, it will **analyze and correlate data** received from cyber, physical and cyber-physical sources, and provide actions that need to be taken in order to alleviate identified issues.
The CPSEC project is structured around the NIST Cyber Security Framework as shown:

Identify - Protect - Detect - Respond - Recover
Security provisioning with Prescriptive Approach

Data driven intelligence....
Three Phases in Prescriptive Analytics

Predictive Analytics

- What will happen?
- When will it happen?
- Why will it happen?

Predictions

- How do we benefit from these predictions?

Decisions

- How will these decisions impact everything else?

Effects
References

• J. Han, M. Kamber, “Data Mining Concepts” Morgan Kaufmann Publications-2002
• R. Agrawal, T. Imielinski, and A. Swami. Mining Association rules between sets of items in large databases. SIGMOD’93.
• J. Han, J. Pei “FP Tree growth” - 2000.
Thank You