



Universiteit
Leiden
The Netherlands

Mining sensor data from complex systems

Vespier, U.

Citation

Vespier, U. (2015, December 15). *Mining sensor data from complex systems*. Retrieved from <https://hdl.handle.net/1887/37027>

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/37027>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/37027> holds various files of this Leiden University dissertation.

Author: Vespier, Ugo

Title: Mining sensor data from complex systems

Issue Date: 2015-12-15

Nederlandse Samenvatting

Tegenwoordig kan vrijwel alles, van natuurverschijnselen tot fysieke systemen, bemeten worden, terwijl de verkregen informatie verzameld, opgeslagen en geanalyseerd wordt om nieuwe inzichten te verschaffen. De inzet van meet-systemen voor allerlei soorten industriële, commerciële en consumententoe-passingen heeft feitelijk verscheidene kansen geschapen voor het analyseren van complexe systemen op een tot op heden ongekend detailniveau.

In dit proefschrift wordt aangetoond dat complexe systemen vaak verschillend gedrag vertonen op verschillende tijdsschalen, en wordt bepleit dat datamining methoden ook op meerdere resoluties (tijdsschalen) zouden moeten werken, om de betreffende data volledig te doorgronden en er zinvolle informatie uit te extraheren.

Onder deze aannames heb ik methods ontwikkeld en geëvalueerd voor datamining en visualisatie van omvangrijke tijdreeksdata over complexe systemen die verzameld wordt door middel van sensoren. Specifiek heb ik oplossingen aangedragen voor drie fundamentele vraagstukken: het detecteren van patronen op meerdere tijdschalen, het herkennen van herhalende gebeurtenissen, en het interactief visualiseren van uitzonderlijke grote tijdreeksdata.

De door mij geïntroduceerde methoden en algoritmen combineren concepten van datamining, signaalverwerking en de informatietheorie. Ik demonstreer hoe verschillende technieken gecombineerd kunnen worden om de uitdagingen aan te pakken die naar voren komen bij het analyseren van tijdreeksdata uit de praktijk. Voorbeelden daarvan zijn de aanwezigheid van ruis in de metingen, het vóórkomen van twijfelachtige en afwijkende gebeurtenissen, en als laatste het risico van *overfitten* van de data met modellen die nauwelijks

nog generaliseren.

Hoewel de behandelde onderzoeksvragen in dit proefschrift een brede toepasbaarheid hebben, worden de voorgestelde oplossingen geëvalueerd op een specifieke toepassing vanuit InfraWatch, een *Structural Health Monitoring* project dat zich richt op het verwerken en analyseren van data geproduceerd door een sensornetwerk op een Nederlandse snelwegbrug.

De door mij ontwikkelde methodes maken het mogelijk om de relevante tijdschalen te detecteren die gelden in de InfraWatch gegevens (alsook in andere databronnen). Daarnaast worden de verschillende herhalende patronen (zogenoemde *motifs*) ontdekt, en wordt de interactieve visualisatie van terabytes aan tijdreeksdata mogelijk gemaakt.