**Institution: City University London**

**Unit of Assessment: 11 Computer Science and Informatics**

**Title of case study: Design diversity for safety and reliability in software-based systems**

## 1. Summary of the impact

Research in the Centre for Software Reliability (CSR) at City University London has made significant advances in ways to assess the safety and reliability of safety-critical, fault-tolerant software-based systems. This work supports quantitative safety cases and has influenced practice and regulation in UK and international industries. [text removed for publication] The work has had significant benefit for regulators and licensees of UK nuclear plant, has been recognised in the US nuclear industry and is additionally of benefit to the general public, in ensuring not only that reasoning about the safety of nuclear plant is rigorous and valid, but also that it is seen to be so in order that safety claims are widely and justifiably believed.

## 2. Underpinning research

It is a truism that systems are becoming increasingly reliant on software for their correct functioning. This poses particular difficulties for safety-critical systems, which often have very stringent reliability requirements and where failure might result in extensive loss of life and significant economic losses. How can we make such systems sufficiently safe? How can we assess a particular system to be confident that it has the required reliability?

An attractive design solution is the use of software diversity in fault tolerant architectures. The use of diversity to make things dependable is age-old (informally: “belt and braces”). In engineered systems, protective redundancy has long been an important approach for achieving high dependability, e.g., the use of replicated stand-by systems. Of course, simple replication does not work for software: multiple copies of software will fail together. Research undertaken in the CSR at City University London instead deals with the use of diversity. Here diversity means “doing things differently”: e.g., software versions that are intellectually different because they have been developed independently by different teams, using different processes.

While this approach has a common sense appeal, there are important technical issues that arise in the use of diversity in fault-tolerant software-based systems such as multi-channel safety protection systems. One of the most important – and the focus of the majority of the CSR research – is the problem of assessment. What we want is a low probability of the fault-tolerant system being defeated by common failures of the diverse redundant components. How can we assess this for real systems? A related research theme concerns methods of diversity achievement: what are the processes and practices that produce diversity most cost-effectively and are beneficial in delivering system dependability?

CSR's work on diversity began in the 1980s and is still thriving. It encompasses both probabilistic modelling and experimental work. The team’s achievements in recent years have extended the scope of knowledge on diversity to cover wider ranges of system types, ways that diversity is pursued and ways that it is threatened. The following are just a few, selected for their impact to date:

1. Our rigorous probabilistic modelling demonstrated that previous claims based on assuming independence of failures between diverse channels (used in safety cases for some real critical systems) are wrong (possibly dangerously so), even when channels are functionally diverse.[1,2]
2. We devised alternative, provably conservative, quantitative (probabilistic) approaches for such safety assessments.[6]
3. We discovered, via rigorous BBN (Bayesian Belief Net) modelling, that there can be subtle and counter-intuitive problems in quantitative safety case reasoning, and we developed ways to avoid some of them.[5]
4. We devised a formal theory of probabilistic “confidence” in system safety claims and showed how it could be applied to diverse *arguments* in support of a claim (so-called multi-legged arguments).[5]

5. We proved that for a special architecture, in which one channel of a 1-out-of-2 system is “possibly perfect”, the simple product of channel A’s pfd (probability of failure on demand) and channel B’s pnp (probability of non-perfection) is a conservative bound on system pfd (in contrast to the usual situation in which each channel must be considered imperfect: here the system pfd is not simply the product of the channel pfds)[6].
6. We devised guidance, based on probabilistic modelling together with empirical research, to aid decision-making by designers of design-diverse systems.[3,4]
7. In a case study of a decision aid computer system for medical use, we obtained novel and surprising results about human/computer diversity. We obtained new insight about limits to the effectiveness of such tools in reducing errors and showed that they might even cause some errors.[7]
8. We have applied these methods to diversity for security with, for example, experimental results on the security increase achievable by combining diverse malware detection tools.

The research team (all employed at City) comprises Professor B Littlewood (1966 to present), Professor L Strigini (1995 to present), Dr P Popov (from 2000, currently Reader), Professor R Bloomfield (2000 to present), Professor P Bishop (2000 to present), Dr A Povyakalo (Research Fellow from 2001, Senior Lecturer from 2008), Dr I Gashi (Lecturer from 2012) and Dr V Stankovic (Lecturer from 2013). Research Fellows employed across the period include Drs Pizza, Bosio, van der Meulen, Gierl and, still present, Alberdi (from 2001), Salako (from 2002) and Wright (from 1986).

## 3. References to the research

[1] Littlewood, B. (1996). The impact of diversity upon common mode failures, *Reliability Engineering and System Safety*, 51(1), 101-113 10.1016/0951-8320(95)00120-4

[2] Littlewood B., Popov P. & Strigini, L. (1999). A note on reliability estimation of functionally diverse systems. *Reliability Engineering and System Safety*, 66(1), 93-95 10.1016/S0951-8320(99)00014-9

[3] Littlewood B., Popov P.T., Strigini L. & Shryane N. (2000). Modeling the Effects of Combining Diverse Software Fault Detection Techniques, *IEEE Transactions on Software Engineering*, *26*, 1157-1167 10.1109/32.888629

[4] Strigini L. & Littlewood, B. (2000). *A discussion of practices for enhancing diversity in software designs*, London: UK, Centre for Software Reliability, City University London, DISPO project Technical Report LS_DI_TR_04 http://openaccess.city.ac.uk/275/

[5] Littlewood B. & Wright D. (2007). The use of multilegged arguments to increase confidence in safety claims for software-based systems: A study based on a BBN analysis of an idealized example. *IEEE Transactions on Software Engineering*, 33(5), 347-365 10.1109/TSE.2007.1002 (This paper was chosen as the “Spotlight Paper” of the issue in which it was published.)

[6] Littlewood, B. & Rushby, J. (2012). Reasoning About The Reliability Of Diverse Two-Channel Systems In Which One Channel Is “Possibly Perfect”, *IEEE Transactions on Software Engineering*, 38(5), 1178-1194 10.1109/TSE.2011.80 (This paper was chosen as the “Spotlight Paper” of the issue in which it was published.)

[7] Alberdi E., Povyakalo A.A., Strigini L., Ayton, P. (2004). Effects of incorrect CAD output on human decision making in mammography, *Academic Radiology*, 11(8), 909-918 10.1016/j.acra.2004.05.012 (recipient of the 2005 Herbert M. Stauffer Award for "Best Clinical Paper" from the Association of University Radiologists), 2004

The selected articles are published in highly-regarded journals which apply rigorous peer review. The technical report was produced for the DIverse Software PrOject (DISPO) which has continued to receive funding from the industry (see below).

The work has received continuous funding for the past 17 years from the UK nuclear industry’s Control & Instrumentation Nuclear Industry Forum (CINIF) (1996 to present), over £1M for DISPO, over £2M in funding from EPSRC across four major projects, £240k from the Leverhulme Foundation and £335k from the European Commission ARTEMIS Joint Technology Initiative.

## 4. Details of the impact

The CSR team has had close involvement with the UK nuclear industry for almost 20 years. The provision of continuous funding for the last 17 years from CINIF attests to the value of our work on diversity. CINIF coordinates research into control and instrumentation in the nuclear industry for the Government Health and Safety Executive's (HSE) Office for Nuclear Regulation (ONR) and comprises all the major nuclear licensees. It has a particular focus on the adequacy of computer-based safety systems.

The industry used diversity in design and implementation long before there were computer systems playing critical safety protection roles. With the introduction of software-based systems, there were attempts to argue that simple notions of independence of failure could be used to support claims for very small system pfd: e.g., for a 1-out-of-2 system, the system pfd could be claimed to be better than $10^{-6}$ if each channel were better than $10^{-3}$. The CSR research showed rigorously that such claims could not be justified, indeed that they were likely to be too optimistic. This allowed, indeed required, regulators to impose stricter requirements on nuclear licensees in making safety cases for multi-channel software-based protection systems.

In January 2008, the Government published a White Paper on the future of nuclear power which concluded that it would be in the public interest to allow energy companies to invest in new nuclear power stations. The ONR with the Environment Agency consequently conducted a Generic Design Assessment (GDA) in relation to the nuclear safety and security aspects of four reactor designs. CSR research results concerning the limitations of what can be claimed for diversity played an important role in the ensuing discussions between regulators and licensees concerning the safety of the protection systems of the proposed new UK reactors.[8]

[text removed for publication]

Complementing the important "negative" results has been extensive work on what might be done in the face of these limitations. For example, we have provided guidance on means to achieve diversity between channels (albeit falling short of guaranteeing failure independence). In addition, from our extensive probabilistic modelling work we have many results that allow system pfd claims to be justified. Essentially these results are provably conservative claims based on assumptions that fall short of independence. For example, in the case of a 1-out-of-2 system in which one channel is "possibly perfect" we have proved that the system pfd is better than the product of channel A's pfd and channel B's pnp[6.] This is the kind of reasoning needed for the EPR (European Pressurised Reactor) safety case, where a simple possibly-perfect third channel is to be added to the originally proposed system.[10] At the CINIF meeting in June 2013 CSR was asked to prepare an extensive technology transfer programme to disseminate the DISPO project results to practising safety engineers; funding approval for this has now been given for it to commence in 2014.

The most obvious beneficiaries of this research are regulators and licensees of UK nuclear plant. Our diversity work has also received recognition in the US nuclear industry. A 2010 report for the US Nuclear Regulatory Commission on the technical basis for establishing acceptable mitigating strategies for nuclear safety systems devotes a section to consideration of the DISPO findings, concluding that "*diversity usage based on the DSDs [diversity-seeking decisions] identified through the UK DISPO research can be considered to provide a very thorough approach to resolving the potential for CCF [common-cause failure] vulnerabilities in software-based Instrumentation & Control systems.*"[11] In addition, of course, there is benefit to the general public, in ensuring not only that reasoning about the safety of nuclear plant is rigorous and valid, but also that it is seen to be so in order that safety claims are widely (and justifiably) believed.

## 5. Sources to corroborate the impact

[8] Bev Littlewood, 'Comments on "Step 3 C&I Assessment of the EDF and AREVA UK EPR" (Division 6 Assessment Report No. AR 09/038-P)', 26 January, 2010. Memorandum in response to invitation to comment by HSE, New Reactor Build, Joint Programme Office. Response to this from

the Office for Nuclear Regulation also available.

[9] A supporting statement (extract quoted above) from ONR's Systems Manager for the Generic Design Assessment of new reactors.

[10] HSE, Office for Nuclear Regulation, ‘Step 4 Control and Instrumentation Assessment of the EDF and AREVA UK EPR Reactor’, ONR-GDA-AR-11-022, 11 November 2011.

[11] NUREG/CR-7007, R. T Wood, R. Belles, et al, ‘Diversity Strategies for Nuclear Power Plant Instrumentation and Control Systems’, US Nuclear Regulatory Commission, 2010. http://pbadupws.nrc.gov/docs/ML1005/ML100541256.pdf

Further information to corroborate claims can be provided by:
ONR (concerning impact on regulation of new reactors)
EdF Energy (operator of U.K. nuclear reactors)
Oak Ridge National Laboratory, US (author of report 11)